

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 16:08:36 ; Search time 4402 Seconds  
(without alignments)  
10570.999 Million cell updates/sec

Title: US-09-830-810A-1  
Perfect score: 1277  
Sequence: 1 aaggcgccgagcgccgga.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl1:\*

9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	643	50.4	675	2	BB641267
2	642.4	50.3	666	2	BB264222
3	603.4	47.3	845	7	CO800921
C	591.4	46.3	615	4	EG071693
5	582.4	45.6	595	4	EG084538
6	571.4	44.7	731	7	CO809618
7	570.4	44.7	594	2	BF471866
8	522.8	40.9	722	7	CO815317
9	502.4	39.3	546	7	CO797395
10	497.2	38.9	525	2	BB704019
11	480.4	37.6	687	7	CO804482
C	473.8	37.1	499	1	AI854700
13	448.4	35.1	491	2	BB703259
14	413.8	32.4	491	2	BB704648
C	405.8	31.8	436	1	AU044294
15	398.2	31.2	425	2	BB705931
C	376.4	31.0	590	1	AU023153
18	393	30.8	450	2	BB946858
19	392.2	30.7	419	2	BB704449
20	367.4	28.8	521	2	BB703869
21	367	28.7	427	2	BB706957
22	364.2	28.5	434	2	BB699732
23	347.2	27.2	391	2	BB704146
24	333.8	26.1	400	2	BB700620

25	319	25.0	905	8	BZ114495
26	316.2	24.8	441	2	BB704848
C	282.2	22.1	301	5	BB637052
28	284.6	20.7	315	2	BB248342
29	227.4	17.8	642	6	BY179704
30	224.8	17.6	654	8	AZ225053
31	211.6	16.6	251	1	AV359580
C	208.6	16.3	251	1	AV358045
33	207.4	16.2	819	5	BP740232
C	203.2	15.9	249	1	AV357766
35	203	15.9	290	2	BB631113
C	201.8	15.8	539	1	AL904756
C	200.8	15.7	909	5	EX755136
38	200	15.7	286	2	BB060381
39	199.8	15.6	246	2	BB060134
C	198	15.5	539	1	AL904892
C	197.8	15.5	642	4	EM861479
C	197	15.4	643	4	BU098465
C	196.8	15.4	526	1	AL904839
C	196.8	15.4	534	1	AL904820
C	196.8	15.4	559	1	AL904774

ALIGNMENTS

RESULT 1  
LOCUS BB641267  
DEFINITION BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus musculus CDNA clone A830014H23 5', mRNA sequence.  
ACCESSION BB641267  
VERSION BB641267.1 GI:16476392  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 675)  
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)





FEATURES  
source

Db	62	CATCTTATCCGAGGCCACCAAGCCGGGATGGCTGGAGGTTCCGACCCAGGCGCTGC	121
Qy	121	CGACCGCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTC	180
Db	122	CGACCGCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTC	181
Qy	181	GACAGCACACGAGCGGCGACAGCTCATGGCCCTGCTGTCGCGGATGGGTCCCCGGT	240
Db	182	GACAGCACACGAGCGGCGACAGCTCATGGCCCTGCTGTCGCGGATGGGTCCCCGGT	241
Qy	241	AGCAGCGGTGACGCTGCGGTGAGGTGAAACCGCGCCCGACACGCTCGGTGCGAGTGTCA	300
Db	242	AGCAGCGGTGACGCTGCGGTGAGGTGAAACCGCGCCCGACACGCTCGGTGCGAGTGTCA	301
Qy	301	CTCGGCGCGCGACGCTGCGAGCTTCGAGGTTGCCGAGCGACGCCCGCATCGGCT	360
Db	302	CTCGGCGCGCGACGCTGCGAGCTTCGAGGTTGCCGAGCGACGCCCGCATCGGCT	361
Qy	361	TCCTGTCAACCCCGTGGCGACGCGCGCGCGCGGAGATCCCGCGATCTCGGACACCGTA	420
Db	362	TCCTGTCAACCCCGTGGCGACGCGCGCGCGCGGAGATCCCGCGATCTCGGACACCGTA	421
Qy	421	GCCCCGTTCTGTTCCGTGACCTTCTGTGGCCCTCTCTCTCTCACTGGAGGTTGCGGAGGC	480
Db	422	GCCCCGTTCTGTTCCGTGACCTTCTGTGGCCCTCTCTCTCTCACTGGAGGTTGCGGAGGC	481
Qy	481	AGGCAGACACCCACGAAGGAGGAGGGGAGCCCGGATCTCTGGGACCCCGGAAACCGGAG	540
Db	482	AGGCAGACACCCACGAAGGAGGAGGGGAGCCCGGATCTCTGGGACCCCGGAAACCGGAG	541
Qy	541	CCGAGAGAGTGGCCCGGAGGAAAGCGGTGCCCGACGCCGCGAAGCGAGGAGGCGATGTT	600
Db	542	CCGAGAGAGTGGCCCGGAGGAAAGCGGTGCCCGACGCCGCGAAGCGAGGAGGCGATGTT	600
Qy	601	CAGGCTGCAGGGCAGCGCGGTTGGGACGA	629
Db	601	CAGGCTGCAGGGC-GGCGGGTGGGACGA	628
RESULT 4			
BG071693/c			
LOCUS			
DEFINITION			
H3102B03-3 NIA Mouse 15K cDNA Clone Set Mus musculus CDNA clone			
H3102B03 3', mRNA sequence.			
ACCESSION			
BG071693			
VERSION			
BG071693.2 GI:40072037			
BG071693			
615 bp mRNA linear EST 18-DEC-2000			

RESULT 4	615 bp	mRNA	linear	EST 18-DEC-2003
BG071693/c				
LOCUS				
DEFINITION	3' NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone			
ACCESSION	H3102B03.3			
VERSION	H3102B03.3			
KEYWORDS	3', mRNA sequence.			
SOURCE	BG071693			
ORGANISM	BG071693.2			
REFERENCE	GI:40072037			
AUTHORS	EST.			
TITLE	Mus musculus (house mouse)			
JOURNAL	Mus musculus			
MEDLINE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
PUBLISHED	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
COMMENT	1 (bases 1 to 615)			
	Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,			
	Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,			
	Wood,W.H. III, Becker,K.G. and Ko,M.S.H.			
	Genome-wide expression profiling of mid-gestation placenta and			
	embryo using a 15,000 mouse developmental cDNA microarray			
	Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)			
	20381348.			
	10922068			
	On Jan 26, 2001 this sequence version replaced gi:12554262.			
	Other ESTs: H3102B03-5			
	Contact: George J. Kargul			
	Laboratory of Genetics			
	National Institute on Aging/National Institutes of Health			
	333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA			
	Email: cdna@lgsun.grc.nia.nih.gov			
	This clone set has been freely distributed to the community. Please			
	visit <a href="http://lgsun.grc.nia.nih.gov/cDNA/15k.html">http://lgsun.grc.nia.nih.gov/cDNA/15k.html</a> for details.			
	Plate: H3102 row: B column: 03			
	Seq primer: -21M13 forward			
	High quality sequence stop: 615			
	POLYA=Yes.			

```

Location/Qualifiers
1. .615
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-3"
/db_xref="taxon:10090"
/clones="H3102B03"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the y chromosome 1998, Hum Mol Genet 7: 1967-1978."

```

## ORIGIN

	Query Match	46.3%;	Score 591.4;	DB 4;	Length 615;	
	Best Local Similarity	99.5%;	Pred. No. 1.7e-131;			
	Matches 614;	Conservative	0;	Mismatches 1;	Indels 2;	Gaps 2;
QY	643	GAGGACCGAA	CAGTGTGCGCGCATG	CGTCTGAGCCCTGGGAGCGAGGACG	CAATGTCT	702
DB	615	GAGGACCGAA	CAGTGTGCGCGCATG	CGTCTGAGCCCTGGGAGCGAGGACG	CAATGTCT	556
QY	703	GCOCGACAGATGGCT	CAGGACCCCGTGTGATTCGGATG	CGCCCTTCGAGACGAGCCCTCCCG		762
DB	555	GCOCGACAGATGGCT	CAGGACCCCGTGTGATTCGGATG	CGCCCTTCGAGACGAGCCCTCCCG		496
QY	763	CRAAGCAGCGAGCAG	GACGAAGAGCGCCTGCTTTCCAGTTC	TCTTAGAGCAGAAAGTACGGC		822
DB	495	CRAAGCAGCGAGCAG	GACGAAGAGCGCCTGCTTTCCAGTTC	TCTTAGAGCAGAAAGTACGGC		436
QY	823	TACTATCACTGCAAG	GACTGCAAAATCCGGTGGGAGCGCCT	TATGTGTGTGTGTCAG		882
DB	435	TACTATCACTGCAAG	GACTGCAAAATCCGGTGGGAGCGCCT	TATGTGTGTGTGTCAG		376
QY	883	GGCACCAAGTAAAGT	GTACTTCAACAGTGTCTGCCAGTGTGTG	GAGAAATCTTACAACC		942
DB	375	GGCACCAAGTAAAGT	GTACTTCAACAGTGTCTGCCAGTGTGTG	GAGAAATCTTACAACC		317
QY	943	TTACAGAGTGGAGCA	CATCACCTGTCAAGTGTGTAAAGAACT	TAGATGTGCCTGCCCAGT		1002
DB	316	TTACAGAGTGGAGCA	CATCACCTGTCAAGTGTGTAAAGAACT	TAGATGTGCCTGCCCAGT		257
QY	1003	CAGATTTTCGCCAC	TGGACCTTAAACGCCCCCATCGC	CAAGACTTGTGTGGGAGATGC	CAA	1062
DB	256	CAGACTTTGCCAC	TGGACCTTAAACGCCCCCATCGC	CAAGACTTGTGTGGGAGATGC	CAA	197
QY	1063	GGACAAACCGCTGT	CTGTGGAGACGACCTTCAGCTTCA	AAATAATCATCATTTAGTGAGAGTC		1122
DB	196	GGACAAACCGCTGT	CTGTGGAGACGACCTTCAGCTTCA	AAATAATCATCATTTAGTGAGAGTC		137
QY	1123	GAAACCGTTTCTGT	AGATGGGGCTAATGAAATGGACA	AGTGAGCTTCTCCCTCTTCA		1182
DB	136	GAAACCGTTTCTGT	AGATGGGGCTAATGAAATGGACA	AGTGAGCTTCTCCCTCTTCA		77

Qy 1183 CCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTGTTACTTGGATATAAAGCCGTGAA 1242  
 Db 76 CCTCTTCCCTTTCCAAATTTCTTCATGACAGACAGTG-TACTTGGATATAAAGCCGTGAA 18

Qy 1243 TAAAGGTATTGCAAAAC 1259  
 Db 17 TAAAGGTATTGCAAAAC 1

RESULT 5  
 BG084538  
 LOCUS  
 DEFINITION H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
 H3102B03 5', mRNA sequence.  
 ACCESSION BG084538  
 VERSION BG084538.2 GI:40072038  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 595)  
 Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,  
 Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,  
 Wood,W.H. III, Becker,K.G. and Ko,M.S.H.  
 Genome-wide expression profiling of mid-gestation placenta and  
 embryo using a 15,000 mouse developmental cDNA microarray  
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
 20381348  
 MEDLINE  
 PUBMED  
 COMMENT On Jan 26, 2001 this sequence version replaced gi:12567102.  
 Other ESTs: H3102B03-3  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lgsun.grc.nia.nih.gov  
 This clone set has been freely distributed to the community. Please  
 visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
 Plate: H3102 row: B column: 03  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 595  
 POLYA=No.

FEATURES  
 source  
 1..595  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:H3102B03-5"  
 /db\_xref="taxon:10090"  
 /clone="H3102B03"  
 /sex="Clones arrayed from a variety of cDNA libraries"  
 /dev\_stages="Clones arrayed from a variety of cDNA libraries"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse 15K cDNA Clone Set"  
 /note="Vector: pSPORT1; Site1: SalI; Site2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryos cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from

## ORIGIN

Query Match 45.6%; Score 582.4; DB 4; Length 595;  
 Best Local Similarity 99.7%; Pred. No. 2.6e-131;  
 Matches 594; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 Qy 511 CCGCATCTCCGGGACCCGGGACCGGACCGGAGAGAGGTGGCGCGGAGAAAGCGGTC 570  
 Db 1 CCGGCATCTCTGGGGACCCGGGAAACCGGAGCCGAGAGAGGTGGCGCGGAGAAAGCGGTC 60  
 Qy 571 CCCAGCGCGCAAGCAGGAGGCGGATGTTTCAGGCTGCAGGCGCAGGCGCGGTGGAGCAG 630  
 Db 61 CCCAGCGCGCAAGCAGGAGGCGGATGTTTCAGGCTGCAGGCGCAGGCGCGGTGGAGCAG 120  
 Qy 631 CAGCCACACCGGAGGACCCGGAACAGTGTGGCGCGATGTCAGTCTGAGCCTGGGAGCGAG 690  
 Db 121 CAGCCACACCGGAGGACCCGGAACAGTGTGGCGCGATGTCAGTCTGAGCCTGGGAGCGAG 180  
 Qy 691 GAGCCATGCTCTGGCGCAGAGATGGCTCAGGACCCCGGTGATTCGGATGCCCTCGAGAC 750  
 Db 181 GAGCCATGCTCTGGCGCAGAGATGGCTCAGGACCCCGGTGATTCGGATGCCCTCGAGAC 240  
 Qy 751 CAGGCTCTCCCGCAAAAGCAGGAGCAGGACAAAGAGGCGCCTGCGTTCAGTCTTAGAG 810  
 Db 241 CAGGCTCTCCCGCAAAAGCAGGAGCAGGACAAAGAGGCGCCTGCGTTCAGTCTTAGAG 300  
 Qy 811 CAGAAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCTATGTG 870  
 Db 301 CAGAAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCTATGTG 360  
 Qy 871 TGGTGTGTCAGGCGCACCACTAGGTGTTACTTCAAACAGTTCGCCGAGTGTGTGAGAA 930  
 Db 361 TGGTGTGTCAGGCGCACCACTAGGTGTTACTTCAAACAGTTCGCCGAGTGTGTGAGAA 419  
 Qy 931 ATCTCTACAACCTTTACAGAGTGGAGGACATCACCTGTCAAAGTTGTAAAGAACTAGATG 990  
 Db 420 ATCTCTACAACCTTTACAGAGTGGAGGACATCACCTGTCAAAGTTGTAAAGAACTAGATG 479  
 Qy 991 TGCCTGCCAGTCAAGATTTGCGCACCTGAGGACCCCTAAACGCCCCCATCGGCAAGACTTGTG 1050  
 Db 480 TGCCTGCCAGTCAAGATTTGCGCACCTGAGGACCCCTAAACGCCCCCATCGGCAAGACTTGTG 539  
 Qy 1051 TGGGAGATGCAAGGACAAACGCCCTGCTCTGCGAGCAGACCTTCAGCTTCAATACA 1106  
 Db 540 TGGGAGATGCAAGGACAAACGCCCTGCTCTGCGAGCAGACCTTCAGCTTCAATACA 595

## RESULT 6

LOCUS CO809618  
 DEFINITION AGENCOURT 30247201 NIH MGC 256 Mus musculus cDNA clone  
 IMAGE:30936113 5', mRNA sequence.  
 ACCESSION CO809618  
 VERSION CO809618.1 GI:51027468  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 731)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.  
 COMMENT Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Kathleen Horner, Stanford University  
 cDNA Library Preparation: Express Genomics

extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."



```
|||||
594 GCGCAGGAAAGCGGTCCCGCCAGCGCGGAAGCGAGGCGCGATGTTCAAGGCTGCAGGGCA 535
QY GCGCGGTGGGAGCAGCAGCACCACCGAGGACCGGACAGTGTGGCGGCGATGCAGTC 674
Db GCGCGGTGGGAGCAGCAGCACCACCGAGGACCGGACAGTGTGGCGGCGATGCAGTC 475
QY TGAGCCTGGGAGCAGGAGCCATGTCCTCCGCGCAGAGATGGCTCAGGACCCCGGTGATTC 734
Db TGAGCCTGGGAGCAGGAGCCATGTCCTCCGCGCAGAGATGGCTCAGGACCCCGGTGATTC 415
QY GGATGCCCTCGAGACAGAGCTCCCGCAAGCAGCGAGCAGGACGAGAGCGCTCGG 794
Db GGATGCCCTCGAGACAGAGCTCCCGCAAGCAGCGAGCAGGACGAGAGCGCTCGG 355
QY TTTCCAGTCTTAGAGCAGAGTAGTACGCTACTATCACTGCAAGGACTCAAAATCCGGTG 854
Db TTTCCAGTCTTAGAGCAGAGTAGTACGCTACTATCACTGCAAGGACTCAAAATCCGGTG 295
QY GGAGAGCGCCTATGTTGTTGTCAGGGCACCAGTAGAGTGTCTTCAAAACAGTTCT 914
Db GGAGAGCGCCTATGTTGTTGTCAGGGCACCAGTAGAGTGTCTTCAAAACAGTTCT 236
QY GCCGAGTGTGAGAAATCTTACAACTTACAGAGTGGAGACATCACTGTCAAAGTT 974
Db GCCGAGTGTGAGAAATCTTACAACTTACAGAGTGGAGACATCACTGTCAAAGTT 176
QY GTAAAGAACTAGATGTCCTGCCAGTCAAGTTCGCGCAGTGGACCCCTAAACGCCCC 1034
Db GTAAAGAACTAGATGTCCTGCCAGTCAAGTTCGCGCAGTGGACCCCTAAACGCCCC 116
QY ATCGCAAGACTTGTGTGGAGATGCAAGGACAAACGCTGTCTTGCACAGCACCTTCA 1094
Db ATCGCAAGACTTGTGTGGAGATGCAAGGACAAACGCTGTCTTGCACAGCACCTTCA 56
QY GCTTCAAAATACATATTTAGTAGAGTGCAGAAACGTTTCTGTAGATGGGCTAAT 1150
Db GCTTCAAAATACATATTTAGTAGAGTGCAGAAAC- TTTCTGTAGATGGGCTAAT 1
```

```
RESULT 8
COB15317 722 bp mRNA linear EST 06-AUG-2004
LOCUS AGNCOURT 30256635 NIH MGC 256 Mus musculus cDNA clone
DEFINITION IMAGE:30939000 5', mRNA sequence.
ACCESSION COB15317
VERSION COB15317.1 GI:51033964
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 722)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1198 row: n column: 01
High quality sequence stop: 495.
Location/Qualifiers
1..722
/organism="Mus musculus"
```

```
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30939000"
/lab_host="DH10B_TonA"
/clone_lib="NIH_MGC_256"
/note="organ: oocyte; Vector: pExpress-1; Site:1: EcoRV;
Site 2: NotI; cDNA was primed using oligo-dr primer:
5'-pGACTAGTTCTAGATCGAGCGCGCC(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb
resulted in an average insert size of 1.2 kb. This is a
primary library (normalized insert size) is NIH_MGC_257)
and was constructed by Express Genomics (Frederick, MD).
Note: this is a NIH_MGC library"
```

## ORIGIN

```
Query Match 40.9%; Score 522.8; DB 7; Length 722;
Best Local Similarity 91.8%; Pred. No. le-116;
Matches 561; Conservative 0; Mismatches 47; Indels 3; Gaps 1;

QY 7 GCGCAGGCGCGGAGCGACCCATGTTCCCGCGAGCAGCTTCCACCCCTGCGCGCATCCT 66
Db |||||
QY 67 TATCCGAGGCGCACAAAGCCGGGGATGGCTGGAGGTTTCGAGCGCAGGGGTGCGGACCC 126
Db |||||
QY 64 TATCCGAGGCGCACAAAGCCGGGGATGGCTGGAGGTTTCGAGCGCAGGGGTGCGGACCC 123
Db |||||
QY 127 GCGCGCCCTCTCTTCCCTCCCGCTACAGACAGCTCATGGCCGCGGATGGTCCCGCTCAGCAGC 246
Db |||||
QY 187 CACGAGGCGGACAGCTCATGGCCGCGGATGGTCCCGCTCAGCAGC 243
Db |||||
QY 247 CGTGACGTGCGGTGCGAGGTGAACCCGCGCGCGAGCTCGTGCGAGTGTTCATCTCGGG 306
Db |||||
QY 244 CGTGACGTGCGGTGCGAGGTGAACCCGCGCGCGAGCTCGTGCGAGTGTTCATCTCGGG 303
Db |||||
QY 307 GCGCGCAGCTGCGAGCGCTGCGAGGTCGCGCAGCGCCCGACGCCCATCGGTTCTGT 366
Db |||||
QY 304 GCGCGCAGCTGCGAGCGCTGCGAGGTCGCGCAGCGCCCGACGCCCATCGGTTCTGT 363
Db |||||
QY 367 CAACCCCGTGCACACGCGCGCGGAGATTCCTCGGATCCTTGCGAGACCGTAGCCCG 426
Db |||||
QY 364 CAACCCCGTGCACACGCGCGCGGAGATTCCTCGGATCCTTGCGAGACCGTAGCCCG 423
Db |||||
QY 427 TTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 486
Db |||||
QY 424 TTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483
Db |||||
QY 487 ACACCCAGAGGAGGAGGAGCGCGCGCATCTCTCGGACCCCGGAGACCGGAGCCGAGA 546
Db |||||
QY 484 ACACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db |||||
QY 547 GAGGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
Db |||||
QY 541 GAGAGTGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db |||||
QY 607 GCAGGGCAGGC 617
Db |||||
601 GCGCGGTGGAC 611
```

## RESULT 9

```
CO797395
LOCUS AGNCOURT 30246265 NIH MGC 257 Mus musculus cDNA clone
DEFINITION IMAGE:30935142 5', mRNA sequence.
ACCESSION CO797395
VERSION CO797395.1 GI:50985575
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
```

CO797395 546 bp mRNA linear EST 05-AUG-2004  
AGNCOURT 30246265 NIH MGC 257 Mus musculus cDNA clone  
IMAGE:30935142 5', mRNA sequence.



ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 546)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. Kathleen Horner, Stanford University  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM1188 row: m column: 07  
High quality sequence stop: 519.  
Features Location/Qualifiers  
1..546  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30935142"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_257"  
/notes="Organ: oocyte; Vector: pExpress-1; Site: 1: EcoRV;  
Site 2: NotI; cDNA was primed using oligo-dT primer;  
5'-pGAGTAGTCTTAGATCGGAGCGGCCCTT25-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb  
resulted in an average insert size of 1.0kb. This is a  
normalized library (primary library is NIH\_MGC\_256) and  
was constructed by Express Genomics (Frederick, MD). Note:  
this is a NIH\_MGC library"

ORIGIN  
Query Match 39.3%; Score 502.4; DB 7; Length 546;  
Best Local Similarity 98.5%; Pred. No. 9.2e-112;  
Matches 525; Conservative 0; Mismatches 6; Indels 2; Gaps 2;  
745 CGAGACAGCGCTCCCGCAAGACGAGCAGGACGAGCAGGAGCGCTGCTTCCAGTTC 804  
1 CGAGACAGCGCTCCCGCAAGACGAGCAGGACGAGCAGGAGCGCTGCTTCCAGTTC 60  
805 TTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGCGCC 864  
61 TTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTGCAAAATCCGGTGGGAGCGCC 120  
865 TATGTGGTGTGTCAGGGACACAGTAAGGTGTACTTCAACAGTTCGCCGAGTGTG 924  
121 TATGTGGTGTGTCAGGGACACAGTAAGGTG-TACTTCAACAGTTCGCCGAGTGTG 179  
925 TGAGAAATCTTACAACTTACAGAGTGGAGGACATCACTGTCAAAAGTGTGTAAGAGAC 984  
180 TGAGAAATCTTACAACTTACAGAGTGGAGGACATCACTGTCAAAAGTGTGTAAGAGAC 239  
985 TAGATGTGCTTCCCGCAGTTCAGATTCGCGACGTCGACCTTAACGCCCTCCATCGGACAGA 1044  
240 TAGATGTGCTTCCCGCAGTTCAGATTCGCGACGTCGACCTTAACGCCCTCCATCGGACAGA 299  
1045 CTTGTGGGAGATGCAAGGCAACCGCTGCTCGGAGCAGCCTTTCAGCTTCAATA 1104  
300 CTTGTGGGAGATGCAAGGCAACCGCTGCTCGGAGCAGCCTTTCAGCTTCAATA 359  
1105 CATCATTTAGTGAGAGTGCAGAAAGCTTCTCTAGATGGGCTTAATGGAATGGCAAGTG 1164  
360 CATCATTTAGTGAGAGTGCAGAAAGCTTCTCTAGATGGGCTTAATGGAATGGCAAGTG 419  
1165 AGCTTTCTCCCTCTTACCTCTTCCCTTCCAAATCTTCATCAGACAGAGTGTACTT 1224

Db 420 AGCTTTCTCCCTCTTCACTCTCCCTTTCCAAATTTCTTCATGACAGACAGTG-TACTT 478  
1225 GGATATAAGCTGTGAATAAAGGATTGCAACCAAAAAAAGAAAAA 1277  
479 GGATATAAAGCTGTGAATAAAGGATTGCAACCAAAAAAAGAAAAA 531

RESULT 10  
BB704019 525 bp mRNA linear EST 11-OCT-2001  
BB704019 RIKEN full-length enriched, in vitro fertilized eggs Mus  
LOCUS musculus cDNA clone 7420459B08 3', mRNA sequence.  
DEFINITION  
BB704019.1 GI:16052854  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
REFERENCE  
AUTHORS  
1 (bases 1 to 525)  
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,  
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akaira, S.,  
Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
TITLE  
JOURNAL  
COMMENT  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: [genome-res@sc.riken.jp](mailto:genome-res@sc.riken.jp), <http://genome.gsc.riken.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.  
Location/Qualifiers  
1..525  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420459B08"  
/sex="female"  
/tissue\_type="in vitro fertilized eggs"  
/dev\_stage="egg"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, in vitro  
fertilized eggs"

FEATURES  
source





QY 548 AGGTGGCCGCGAG 560

Db 537 GTGCCCGAGAAAG 549

RESULT 12

AI854700/c

LOCUS

DEFINITION

UI-M-BH0-akc-d-12-0-UI-s1 NIH BMAP M.S1 Mus musculus cDNA clone

UI-M-BH0-akc-d-12-0-UI 3', mRNA sequence.

ACCESSION

AI854700

VERSION

AI854700.1 GI:5498606

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 499)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643 USA

Tel: 301 443 1706

Fax: 301 443 9690

Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA Library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA=yes.

Location/Qualifiers

1..499

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-BH0-akc-d-12-0-UI"

/dev\_stage="27-32 days"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NIH BMAP M.S1"

/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH BMAP M.S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain.

TAG TISSUE=basal-ganglia

TAG LIB=NIH BMAP M\_S1

TAG\_SEQ=GTGAC"

ORIGIN

Query Match 37.1%; Score 473.8; DB 1; Length 499;

Best Local Similarity 99.2%; Pred. No. 8.9e-105;

Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

772 GAGCAGCAAGGAGCGCGCTTCCAGTTCTTAGAGCAGAGTACGGCTACTATCAC 831

|||||

499 GAGCAGCAAGGAGCGCGCTTCCAGTTCTTAGAGCAGAGTACGGCTACTATCAC 440

|||||

832 TGCAAGGACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGAGGACACAGT 891

|||||

439 TGCNAGGACTGCAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGAGGACACAGT 380

|||||

892 AAGGTGTTACTTCAACAGTTCTGCCGAGTGTGTGAGAAATCTTACNACCTTACAGAGT 951

|||||

379 AAGGTG-TACTTCAACAGTTCTGCCGAGTGTGTGAGAAATCTTACNACCTTACAGAGT 321

|||||

952 GGAGGACATCACCTGTCAAAAGTTGTAAGAAGTCTAGATGTGCTGCCAGCTCAGATTTCG 1011

|||||

320 GGAGGACATCACCTGTCAAAAGTTGTAAGAAGTCTAGATGTGCTGCCAGCTCAGACTTCG 261

|||||

1012 CCAGTGGAGCCCTAAACGCGCCCATCGCAAGCTTGTGTGGGAGATGCAAGCAACG 1071

|||||

260 CCAGTGGAGCCCTAAACGCGCCCATCGCAAGCTTGTGTGGGAGATGCAAGCAACG 201

|||||

1072 CCTGTCTTGGCAGACGACCTTCAAGTTCATCATCATCATCATCATCATCATCATCATCAT 1131

|||||

200 CCTGTCTTGGCAGACGACCTTCAAGTTCATCATCATCATCATCATCATCATCATCATCAT 141

|||||

1132 TCTGTAGATGGGCTTAATGGAATGGAAGTGGAGCTTTCTCCCTCTTCACTCTTCCC 1191

|||||

140 TCTGTAGATGGGCTTAATGGAATGGAAGTGGAGCTTTCTCCCTCTTCACTCTTCCC 81

|||||

1192 TTTCCTCAATTTCTTATGACAGACAGTGTACTTGGATATTAAGCTGTGATATAAGGTA 1251

|||||

80 TTTCCTCAATTTCTTATGACAGACAGTGTACTTGGATATTAAGCTGTGATATAAGGTA 22

|||||

1252 TTGCAACACAAAAA 1272

|||||

21 TTGCAACACAAAAA 1

|||||

RESULT 13

BB703259

LOCUS

DEFINITION

musculus cDNA clone 7420449J15 3', mRNA sequence.

ACCESSION

BB703259

VERSION

BB703259.1 GI:16052094

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 491)

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yashinashi, A., Muramatsu, M. and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

Wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

#### FEATURES

Location/Qualifiers  
1. .491  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="7420449J15"  
/sex="female"  
/tissue type="in vitro fertilized eggs"  
/dev stage="egg"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, in vitro fertilized eggs"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGATCTCGAGTTAATAAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

Query Match 35.1%; Score 448.4; DB 2; Length 491;  
Best Local Similarity 98.2%; Pred. No. 1.4e-98;  
Matches 485; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

Qy 763 CAAGACGGAGCAGGACAGAGCGCTCGGTTTCAGTTCTTAGACGACGAAGTACGGC 822  
Db 1 CAAGACGGAGCAGGACAGAGCGCTCGGTTTCAGTTCTTAGACGAGTACGGC 60

Qy 823 TACTATCACTCAAGGACTGCAAAATCCGGTGGGAGCGCTAGTGTGGTGTGCAG 882  
Db 61 TACTATCACTGCAA-GACTGCAAAATCCGGTGGGAGCGCTAGTGTGGTGTGCAG 119

Qy 883 GGCACCAAGTAGGTGTACTTCAACAGTTCTGCCGAGTGTGTGAGAAATCCTACACCC 942  
Db 120 GGCACCAAGTAGGTG-TACTTTAAACAGTTCTGCCGAGTGTGTGAGAAATCCTACACCC 178

Qy 943 TTACAGAGTGAGGACATCACTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCAGT 1002  
Db 179 TTACAGAGTGAGGACATCACTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCAGT 238

Qy 1003 CAGATTTCCGACGTTGGACCTTAACGCCCCATCGCGCAAGACTTGTGTGGGATGCAA 1062  
Db 239 CAGATTTCCGACGTTGGACCTTAACGCCCCATCGCGCAAGACTTGTGTGGGATGCAA 298

Qy 1063 GGACAAACGGCTGCTCGGACAGCACCTTCAGCTTCAATACATCATTTAGTGAGAGTC 1122

#### ORIGIN

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.  
Location/Qualifiers  
1. .491  
/organism="Mus musculus"  
/mol\_type="mRNA"

#### FEATURES

Location/Qualifiers  
1. .491  
/organism="Mus musculus"  
/mol\_type="mRNA"

Db 299 GGACAAAGCGCTGCTCGGACAGACACCTTCAGCTTCAATACATCATTTAGTGAGAGTC 358  
Qy 1123 GAAACGCTTCTGTAGATGGGGCTAATGGAATGGAACGAAGTGTCTTCCCTCTTCA 1182  
Db 359 GAAACGCTTCTGTAGATGGGGCTAATGGAATGGAACGAAGTGTCTTCCCTCTTCA 418  
Qy 1183 CCTCTCCCTTCCAAATTTCTTCATGACAGACAGTGTCTTGGATATAAAGCCCTGTGAA 1242  
Db 419 CCTCTCCCTTCCAAATTTCTTCATGACAGACAGTGTCTTGGATATAAAGCCCTGTGAA 477  
Qy 1243 TAAAGGTATTGCA 1256  
Db 478 TAAAGGTATTGAA 491

#### RESULT 14

BB704648  
LOCUS BB704648 RIKEN full-length enriched, in vitro fertilized eggs Mus  
DEFINITION musculus cDNA clone 7420466L07 3', mRNA sequence.  
ACCESSION BB704648  
VERSION BB704648.1 GI:16053483  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 491)  
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, K., Shibata, K., Sakai, C., Sakai, K., Sakazume, N., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)  
JOURNAL Unpublished (2001)  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clonename="7420466L07"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clonelib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGACAGGATCCACAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by the
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATCTCCAGGTAAATTAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

```

Query Match	32.4%	Score 413.8	DB 2	Length 491
Best Local Similarity	93.0%	Pred. No. 4.1e-90	Indels 2	Gaps 2
Matches 455	Conservative 0	Mismatches 32		
QY 768	CACGGAGCAGCAAGGAGCGCTGCGTTCCTCCAGTTCCTTAGACAGCAAGTACGGCTACTA	827		
DB 1	CACGGAGCAGCAAGGAGCTCTGCGTTCCTCCAGTTCCTTAGACAGCAAGTACGGCTACTA	60		
QY 828	TCACCTGCAAGGACTGCAAAATCCGGTGGGAGAGCGCCCTATGTGGTGTGTGTCAGGGCAC	887		
DB 61	TCA-AGCAAGGACTGCAAAATCCGCTGGGAAACGCCCTATGTCTGGAGTGTGCGGGCAC	119		
QY 888	CAGTAAGGTGTACTTCAAAACAGTTCGCGAGTGTGAGAAATCTCAACACCCCTTACA	947		
DB 120	CAGTAAGATGTACTTCAAAACAGTTCGCGAGTGTGAGAAATCTCAACACCCCTTACA	179		
QY 948	GAGTGGAGGACATCACCTGTCAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAAGT	1007		
DB 180	GAATGGAGGACATCACCTCTCAAATTTGTAAGAACTAAATGTGCTCTCCAGTCAAGC	239		
QY 1008	TTCCGACAGTGGACCTTAACGCCGCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACA	1067		
DB 240	TTCCGACAGTGGACCTTAACGCCGCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACA	299		
QY 1068	AACGCCCTGCTCCGCGACAGCACCTTCACGCTTCAAATACATCATTTAGTGAGAGTCGAAA	1127		
DB 300	AACGCCCTGCTCCGCGACAGCACCTTCACGCTTCAAATACATCATTTAGTGAGAGTCGAAA	359		
QY 1128	CGTTTCTGCTAGATGGGCTTAATGGAATGGCAAGTGAAGCTTCTCCCTCTTCACTCT	1187		
DB 360	CGTTTCTGCTAGATGGGCTTAATGGAATGGCAAGTGAAGCTTCTCCCTCTTCACTCT	419		
QY 1188	TCCCTTTTCCAAATCTTCATGACAGACAGTGTACTTTGGATATATAAGCTCTGTAATAAA	1247		
DB 420	TCCCTTTTCCAAATCTTCATGACAGACAGTGTACTTTGGATATATAAGCTCTGTAATAAA	478		
QY 1248	GGTATTGCA 1256			
DB 479	GGTATTGTA 487			

RESULT 15	436 bp	linear	EST 09-DEC-1998
AU044294/c		mouse sixteen-cell-embryo	
LOCUS	AU044294	cdna Mus musculus	
DEFINITION	AU044294	Mouse sixteen-cell-embryo	
		cdna	
		sequence.	
		JO917G09 3'	

ACCESSION	AU044294
VERSION	AU044294.1 GI:3979844
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 436) Ko,M.S.H., Kitchen,J.R., Wang,X., Threat,T.A., Sun,T., DePalma,G.E., Liang,Y., Kargul,G.J., Sharara,R., Lim,M.K. and Doi,H.
TITLE	Systematic analyses of genes expressed in 16-cell mouse embryo (The ERATO/Doi project at Wayne State University)
JOURNAL	Unpublished (1998)
COMMENT	Contact: Hirofumi Doi Doi Biosymmetry Project, ERATO Japan Science and Technology Corporation (JST) WEG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan Email: hd@bioa.jst.go.jp.
FEATURES	Location/Qualifiers
source	1..436
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="J0917G09"
	/dev_stage="sixteen-cell-embryo"
	/clone_lib="Mouse sixteen-cell-embryo cDNA"
ORIGIN	
Query Match	31.8%; Score 405.8; DB 1; Length 436;
Best Local Similarity	97.0%; Pred. No. 3.6e-88;
Matches 424; Conservative	0; Mismatches 12; Indels 1; Gaps 1;
QY	823 TACTATCACTGCAGGAGTCCAAATCCGGTGGGAGAGCGCCTATGTGGTGTCGAG 882
Db	436 TATTATCACTGCAGGAGTCCAAATCCGGTGGGAGAGCGCCTATGTGGTGTCGAG 377
QY	883 GGCCACCAGTAAGGTGTACTTCAAACGTTCTCCGAGTGTGAGAATACTTACAACCC 942
Db	376 GGCCACCAGTAAGGTGTACTTCAAACGTTCTCCGAGTGTGAGAATACTTACAACCC 317
QY	943 TTACAGATGGAGACATCACTGTCAAAGTGTGTAAGAATCTAGATGTGCCCTGCAGT 1002
Db	316 TTACAGATGGAGACATCACTGTCAAAGTGTGTAAGAATCTAGATGTGCCCTGCAGT 257
QY	1003 CAGATTTTCGCCACGTGGACCTTAACGCCGCCCATCGCAAGATTGTGTGGGAGATGCAG 1062
Db	256 CAGACTTCGCCACGTGGACCTTAACGCCGCCCATCGCAAGATTGTGTGGGAGATGCAG 197
QY	1063 GGACAAACGCGTGTCTGGACAGCACCTTCAGCTTCAAATACATCATCTTAGTGAGAGTC 1122
Db	196 GGACAAACGCGTGTCTGGACAGCACCTTCAGCTTCAAATACATCATCTTAGTGAGAGTC 137
QY	1123 GAAAACGTTTCTGTAGATGGGGCTAATGAATGGACAGTGAAGTCTTCTCCCCTCTTCA 1182
Db	136 GAAAACGTTTCTGTAGATGGGGCTAATGAATGGACAGTGAAGTCTTCTCCCCTCTTCA 77
QY	1183 CCTCTTCCCTTTCCAAATTCITTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAA 1242
Db	76 CCTCTTCCCTTTCCAAATTCITTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAA 18
QY	1243 TAAAAAGGTATTGCAAC 1259
Db	17 TAAAAAGGTATTGCAAC 1

Search completed: January 5, 2005, 19:20:23  
Job time : 4411 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 14:55:31 ; Search time 5526 Seconds  
(without alignments)  
10928.148 Million cell updates/sec

Title: US-09-830-810A-1  
Perfect score: 1277  
Sequence: 1 aagcgcgagcgagcgaggga.....acaaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245.4	97.5	1260	10 AY191415	Mus musculus
2	962.2	75.3	1280	10 AY283175	Rattus no
3	799.4	62.6	3987	10 AY193889	Mus musculus
4	799.4	62.6	144954	2 AC122733	Mus musculus
5	799.4	62.6	200535	10 AL671880	Mouse DNA
6	689.8	54.0	193274	2 AC107686	Mus musculus
7	628	49.2	235390	2 AC126519	Rattus no
8	628	49.2	237695	2 AC125993	Rattus no
9	628	49.2	262139	2 AC127083	Rattus no
10	578.8	45.3	220894	2 AC108848	Mus musculus
11	352	27.6	1275	9 AY191416	Homo sapi
12	239.2	18.7	366	6 CQ716345	Sequence
13	205.8	16.1	1052	5 AY283176	Xenopus 1
14	194.8	15.3	963	5 AY283177	Takifugu
15	194.6	15.2	1084	5 AY283178	Danio rer
16	147.6	11.6	138890	9 AC007970	Homo sapi
17	147.6	11.6	151961	2 AC037443	Homo sapi
18	134.6	11.6	204268	2 AC092163	Homo sapi
19	134.6	10.5	4138	9 AY193890	Homo sapi

20	134.6	10.5	187578	9	AC096952	Homo sapi
21	97.6	7.6	945	5	BX931269	Gallus ga
22	90.6	7.1	187278	5	BX571888	Danio rer
23	90.6	7.1	249791	5	BX537133	Zebrafish
24	73.2	5.7	694	9	BSA335296	
25	72.2	5.7	125020	9	AF429315	Homo sapi
26	64.8	5.1	144000	9	AF429315	Homo sapi
27	63.2	4.9	125020	9	AF429315	Homo sapi
28	62.6	4.9	1393	11	PM11H12G	
29	62	4.9	136551	2	AC048354	Homo sapi
30	60.8	4.8	160327	4	AC149763	Bos taurus
31	60.8	4.8	240521	2	AC149759	Bos taurus
32	60.4	4.7	155337	2	AC116408	Mus muscu
33	58.6	4.6	155337	2	AC116408	Mus muscu
34	58.6	4.6	202851	10	AC115880	Mus muscu
35	58.4	4.6	248550	1	SC0939120	Streptomy
36	58.2	4.6	94058	2	AC141024	Rattus no
37	58	4.5	92310	2	AC108380	Pan trogl
38	57.8	4.5	2000	6	AX655393	Sequence
39	57.2	4.5	143406	10	AC092857	Rattus no
40	57.2	4.5	230329	2	AC142070	Rattus no
41	56.6	4.4	245210	2	AC137771	Homo sapi
42	56.2	4.4	326244	2	AC118627	Mus muscu
43	56	4.4	142402	2	CR407549	Danio rer
44	56	4.4	247644	5	BX470223	Zebrafish
45	56	4.4	250011	5	BX548044	Zebrafish

#### ALIGNMENTS

RESULT 1	AY191415	1260 bp	mrna	linear	ROD 12-MAR-2003
LOCUS	AY191415	Mus musculus	zygote arrest 1 (Zarl)	mrna	complete cds.
DEFINITION	AY191415				
ACCESSION	AY191415				
VERSION	AY191415.1	GI:27808689			
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 1260)				
AUTHORS	Wu, X., Viveiros, M.M., Eppig, J.J., Bai, Y., Fitzpatrick, S.L. and Matzuk, M.M.				
TITLE	Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for the oocyte-to-embryo transition				
JOURNAL	Nat. Genet. 33 (2), 187-191 (2003)				
MEDLINE	22447938				
PUBMED	12539046				
REFERENCE	2 (bases 1 to 1260)				
AUTHORS	Wu, X., Wang, P. and Matzuk, M.M.				
TITLE	Direct Submission				
JOURNAL	Submitted (04-DEC-2002) Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
FEATURES	Location/Qualifiers				
source	1..1260				
	/organism="Mus musculus"				
	/mol_type="mrna"				
	/db_xref="taxon:10090"				
	/map="between DSBUc48 and Ttk"				
gene	1..1260				
	/gene="Zarl"				
5'UTR	1..27				
	/gene="Zarl"				
CDS	28..1113				
	/gene="Zarl"				
	/note="ZARL; maternal factor; contains atypical C8 PHD motif at C-terminus"				
	/codon_start=1				
	/product="zygote arrest 1"				

/protein\_id="AA024706.1"  
 /db\_xref="GI:27808690"  
 /translation="MPASTFHPCHPHYPOATKAGDGRFARGCRPAPPSFLFGRYQ  
 LMAAEYDSHQRAQLWALLSRMGPRSVSRRDAAVQVNPDRDASVQCSLGRRTLPQAGC  
 RAPSDARSGCQPRGHGAGCRSPRWQTVAPFSVTFGLSSSLEVAGGRQPTTKCEG  
 SPASSTGREPREVAARKVAQPPQRESEGDVAAQAGWQQPPFDNRNSVAAQMSEF  
 GSSEPCPAAMADQPGSDAPQASQSTEQKERLRFQLEQKGYGYHCKDCKINR  
 ESAVMVCQGTSKYVQKFCRCVEKSNVRYVEDITCQSKCRTRCAPVRLRHVDPRR  
 PHRQDLGCRKDKRLSCDSTFSFKYII"  
 .1114..1260  
 /gene="Zar1"  
 3'UTR

[illegible]

Qy	843	CAAAATCCGGTGGGAGAGCGCCCTATGTGTGGTGTGTGTGTCAGGGCACCAAGTAAGGTGTACT	902
Db	843	CAAAATCCGGTGGGAGAGCGCCCTATGTGTGGTGTGTGTGTCAGGGCACCAAGTAAGGTG-TACT	901
Qy	903	TCAAAACAGTTCTGCCAGAGTGTGTGAGAAATCCTTACAACCCCTTACAGAGTGGAGGACATCA	962
Db	902	TCAAAACAGTTCTGCCAGAGTGTGTGAGAAATCCTTACAACCCCTTACAGAGTGGAGGACATCA	961
Qy	963	CTGTGTCAAAGTGTCTAAAGAACTAGATGTGCTTGGCCAGTCAAGATTCGCGCCAGTGGACC	1022
Db	962	CTGTGTCAAAGTGTCTAAAGAACTAGATGTGCTTGGCCAGTCAAGATTCGCGCCAGTGGACC	1021
Qy	1023	CTAAACGCCGCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACGCCCTGTCTCTGG	1082
Db	1022	CTAAACGCCGCCCATCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACGCCCTGTCTCTGG	1081
Qy	1083	ACAGCACCTTCAGCTTCAATATACATCATTTAGTGAGAGTGTGAAACGTTTCTCTCTAGATG	1142
Db	1082	ACAGCACCTTCAGCTTCAATATACATCATTTAGTGAGAGTGTGAAACGTTTCTCTCTAGATG	1141
Qy	1143	GGGCTAATGGAATGGCAAGTGAAGTCTTCTCCCTCTTCTCACTTCTCCCTTCCAAATTC	1202
Db	1142	GGGCTAATGGAATGGCAAGTGAAGTCTTCTCCCTCTTCTCACTTCTCCCTTCCAAATTC	1201
Qy	1203	TTCATGACAGACAGTGTACTCTGGATATAAAGCCTGTGAATAAAGGTATTGCAAAACA	1261
Db	1202	TTCATGACAGACAGTGTACTCTGGATATAAAGCCTGTGAATAAAGGTATTGCAAAACA	1260
RESULT 2			
LOCUS	AY283175	1280 bp	linear
DEFINITION	Rattus norvegicus zygote arrest 1 (Zari) mRNA, complete cds.		
ACCESSION	AY283175		
VERSION	AY283175.1	GI:30908932	
KEYWORDS	Rattus norvegicus (Norway rat)		
SOURCE	Rattus norvegicus		
ORGANISM	Rattus norvegicus		
REFERENCE	Wu,X., Wang,P., Brown,C.A., Zilinski,C.A. and Matzuk,M.M.		
AUTHORS	Zygote arrest 1 (Zari) is an evolutionarily conserved gene		
TITLE	expressed in vertebrate ovaries		
JOURNAL	Biol. Reprod. 69 (3), 861-867 (2003)		
MEDLINE	22811438		
PUBMED	12773403		
REFERENCE	2 (bases 1 to 1280)		
AUTHORS	Wu,X. and Matzuk,M.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-APR-2003) Pathology, Baylor College of Medicine, One		
	Baylor Plaza, Houston, TX 77030, USA		

FEATURES source

gene

CDS

```

/codon_start=1
/product="zygote arrest 1"
/contig_id="X00270.7"

```

```

/db_xref="GI:30908933"
/translation="MFPAATGCPHPHYPTAAKGLDGFPGKGRGCRPPSPPLGVGR
OLMAEYDYSYQALMALLSRMPGVRSDAAVQVNPRRDASVCSLGRRTLQPR
RRAEPDAPGSCQSPRACPGPPRSWTVALYSPVTFGLASSLBAVDGRQPTXGE
GRPAFTRETPGEGVAVMKAPVQPSSEGVQVAGQDQEQEPREDPDSVAAQSEIP
GSEPPPPVMAQDPSVAAVDRASPOSTQEKRLRFVLEQKYQGHCKDCNINW

```

Qy	3	GGCGGCGAGCGCGGGACGCA	CCATGTTCCCGCGGAGCAGTTC	CACCCCTGCCCGCA	62
Db	3	GGCGGCGAGCGCGGGACGCA	CCATGTTCCCGCGGAGCAGTTC	CACCCCTGCCCGCA	62
Qy	63	TCCTTATCCGCG	--AGGCCAACAAAGCCGGGATGGCT	TGGAGGTTCCGAGGCCAGGGCGTG	119
Db	63	TCCTTATCCGCGCCACGGGACG	CAAAAGCCGGGATGGCT	TGGAGGTTTGAGAGCCAGGGCGTG	122
Qy	120	CCGACCCGCGCCCCCTCT	TCTTCCCTCCCGGCTACAGACAGCT	CATGCGCGCGGAGTACGT	179
Db	123	CAGGCCGAGCCCCCTCT	TCTTCCCTCCCGGCTACAGACAGCT	CATGCGCGCGGAGTACTT	182
Qy	180	CGACAGCCACGACGGGCGACAG	CTCATGCGCTCTGTGCGGATGGT	TCCCGGTCCGT	239
Db	183	TGACAGCTATCAGCAGCGAGCT	CATGGCTTGTCTGCGAATGGT	TCCCGGCGGT	242
Qy	240	CAGCAGCGGTGACGCTCGGT	TGACAGTGAACCCGCGCGAGACGCT	TCGTCAGATGTC	299
Db	243	CAGCAGCGCGACGCTCGGT	TGACAGTGAACCCGCGCGGATGCCT	TCGGTGCAGTGTTC	302
Qy	300	ACTCGGCGCGCACGCTG	CAGCTGACGGTGCAGGCCAGCC	CGACCCCGATCGGG	359
Db	303	GCTCGGCGCGCACATG	CAGCTGACGCGCGGAGCCAGCC	CGACCCCGGCGCTGG	362
Qy	360	TTCTGTCAACCGGTGGCCACG	CGCGCGCGGGAGATCCCGGAT	TCCTGGCAGACCGT	419
Db	363	TTCTGTCAACCCGCGACCG	CGCGCGGGAGACCCCGGAT	TCCTGGGCGACCGT	422
Qy	420	AGCCCCGTCTCTGTCGTG	ACCTTCTGTGGCTCTCTCTCTCA	CTGAGAGTTGCGGAGG	479
Db	423	CGCCCTGTACTCGCCCTG	ACCTTCTGGTGGCTCTCTCTCTG	CTGAGAGTTGCGGGGA	482
Qy	480	CAGCGACACCCACGAAGG	AGGGAGCGCGGCATCTCTCGGG	ACCCGGGAACCGGA	539
Db	483	CAGCGACACCCACGAAGG	AGGGAGACGGCACCCACCGGG	ACCCGGGAACCCGA	542
Qy	540	GCCGAGAGAGTGGCCCGAG	GAACGGTCCCCCAGCGGAGCG	AGCGAGAGGCGCATGT	599
Db	543	GCCGGGAGAGTGACGTG	ATGAAGCAGTCCCCCAGCGAG	CGAGAGGCGCGCATGT	602
Qy	600	TCAGGCTGACAGGCGAG	CGGGTGGAGCAGCACCCAC	CCGAGAGCCGGAACAGTGT	659
Db	603	CCAGGCTGAGGGCAGGATG	--GGCAGGAGCAGCCAC	CGGGGAGGACCCGAGCATGT	659
Qy	660	GGCGCGCATGCAGTCT	GAGCTGGAGCGAGAGCCATGT	CTGCGCGCAGAGATGGTCA	719
Db	660	GGCGCGCATGCAGTCT	GAGCCGCGGAGTGGAGGCC	ACCTCTGCTGCGATGGTCA	719
Qy	720	GGACCCCGGTGATTTC	GAATGCGCTCCAGACACGAG	CGCTCCCGCAAGCAGGAGGAG	779
Db	720	GGACCCCGGTGACGTG	GCTGCTAGAGACCGGGCT	TCCCCACAGAGCACTGAGAGGA	779
Qy	780	CAAGGAGCGCTGCGTT	TCCAGTTCTTAGAGCAG	AAGTACGGCTACTATCACTGC	839
Db	780	CAAGGAGCGCTGCGTT	TCCAGTTCTTAGAGCAG	AAGTACGGCTACTATCACTGC	839
Qy	840	CTGCAAAATCCGGTGG	GAGACGGCTATGTGGTGTG	TCAGGGCAACAGTAAGGTGT	899
Db	840	CTGCAACATCCGGTGG	GAGACGGCTATGTGGTGTG	TCAGGGCAACAGTAAGGTGT	898
Qy	900	ACTTTCAAAAGTTC	TCGCGGAGTGTGCAGAAA	TCCTACAAACCTTACAGAGTGC	959
Db	899	ACTTTCAAAAGTTC	TGCGCGGTGTGAGAA	GTCTTAAACCCATACCGAGTGC	958
Qy	960	TCACCTGTCA	AAGTGTGTAAGAAAC	TACTAGATGTGCTCGCCAGTGC	1019



<div><div>/product="zygote arrest 1"</div><div>/protein_id="AAC24708.1"</div><div>/db_xref="GI:27808694"</div><div>/translation="WFPASTHPCPHYPQATQKAGDWRFGAGCRPAPSFILPGVYQ LMAEYVDSHORAOI.MALLSRMPSRVSDAAVQVNPRRDASVQCSLGRRTLQPAGC RASPDARSQCPRGHAGAGRSRQWTVAFSSVTCGLSSLSLEVAGRGRTPTKGEG SPASGTRSPERPREVAARNAKVPQPSREGSDVQAAGQAGWEQPPEDPDRNSVAAQSE GSEEPCAAEMAQDPDSDAKPDQASPOSTQDKERLARFPQLEQKYGVYHCKDKLIRW ESAYVWCVGTSKVYKPFQRCVCEKSYNPIRVEDITQSCSKRTKRCAPVLRHVDPKR PHRQDLGRCRCKRLSCDSSTFSFKYII"</div><div>2574..2666</div><div>/gene="Zarl1"</div><div>/number=2</div><div>2745..2819</div><div>/gene="Zarl1"</div><div>/number=3</div><div>3697..3987</div><div>/gene="Zarl1"</div><div>/number=4</div></div>									
exon									
exon									
exon									
ORIGIN									
Query Match 62.6%; Score 799.4; DB 10; Length 3987;									
Best Local Similarity 99.9%; Pred. No. 1e-155; Indels 0; Gaps 0;									
Matches 800; Conservative 0; Mismatches 1;									
QY	3	GGCGGGGAGCGCGGGAGCAGCACCATGTTCCCGCGAGCAGCTTCCACCCCTGCGCGCA	62						
DB	3	GGCGGGGAGCGCGGGAGCAGCACCATGTTCCCGCGGAGCAGCTTCCACCCCTGCGCGCA	62						
QY	63	TCCTTTATCCGAGGCCACCAAGCCGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	122						
DB	63	TCCTTTATCCGAGGCCACCAAGCCGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	122						
QY	123	ACCGCGGCCCCCTCCTTCTCCCGCGCTACAGACAGCTCATGGCGCGGAGTACGTGGA	182						
DB	123	ACCGCGGCCCCCTCCTTCTCCCGCGCTACAGACAGCTCATGGCGCGGAGTACGTGGA	182						
QY	183	GAGCCACAGCGGGCAGAGCTCATGGCCCTGCTGTGCGGATGGGTCCCGCTCGGTGAG	242						
DB	183	GAGCCACAGCGGGCAGAGCTCATGGCCCTGCTGTGCGGATGGGTCCCGCTCGGTGAG	242						
QY	243	CAGCGGTGACGCTGCGGTGACAGTGAAACCGCGCGCGCAGCGCTCGGTGCGAGTTCAT	302						
DB	243	CAGCGGTGACGCTGCGGTGACAGTGAAACCGCGCGCGCAGCGCTCGGTGCGAGTTCAT	302						
QY	303	CGGGCGCGCAGCGCTGACGCTGACAGGTCGAGCCCGCCGAGCCCGCGATCCGGTTC	362						
DB	303	CGGGCGCGCAGCGCTGACGCTGACAGGTCGAGCCCGCCGAGCCCGCGATCCGGTTC	362						
QY	363	CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC	422						
DB	363	CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC	422						
QY	423	CCCGTTCTGTCCTGTGAACCTTCTGTGGCTCTCTCTCTCATCGAGGTTGCGGAGGAG	482						
DB	423	CCCGTTCTGTCCTGTGAACCTTCTGTGGCTCTCTCTCTCATCGAGGTTGCGGAGGAG	482						
QY	483	GAGACACCCAGAGGAGGAGGGAGCCCGGCATCTCTGGGGACCCCGGAAACCGGAGCC	542						
DB	483	GAGACACCCAGAGGAGGAGGGAGCCCGGCATCTCTGGGGACCCCGGAAACCGGAGCC	542						
QY	543	GAGAGAGTGGCCGCGAGAAAGCGGTTCCCTCAGCCCGGAAGCGAGAGGGCGATCTTCA	602						
DB	543	GAGAGAGTGGCCGCGAGAAAGCGGTTCCCTCAGCCCGGAAGCGAGAGGGCGATCTTCA	602						
QY	603	GGCTGCAGGGCAGGGCCGGTGGGACGACGACACACCGGAGGACCGGAACAGTGTGGC	662						
DB	603	GGCTGCAGGGCAGGGCCGGTGGGACGACGACACACCGGAGGACCGGAACAGTGTGGC	662						
QY	663	GGCGATCCAGTCTGAGCCTGGGAGCGAGAGCCATCTCTCCCGCAGATGGCTCAGGA	722						
DB	663	GGCGATCCAGTCTGAGCCTGGGAGCGAGAGCCATGTCTCTCCCGCAGATGGCTCAGGA	722						

[illegible]









Db 27289 TGACAGCTATCAGCGAGCGCAGCTCATGCGCTTGTGTCGGAATGAGTGGTCCCGCGCGGT 27230  
 Qy 240 CAGCAGCGCTGACCTCGGTCGAGGTGAACCCCGCGCGCGCACCTCGGTGCGAGTTC 299  
 Db 27229 CAGCAGCGCGCAGCTCGGTCGAGGTGAACCCCGCGCGCGATCCCTCGGTGCGAGTTC 27170  
 Qy 300 ACTCGGCGCGCAGCTGCGAGCTGCGAGGTGCGCGAGCGCGCGCGCGCGCGCGCGCGG 359  
 Db 27169 GCTCGGCGCGCAGCTGCGAGCTGCGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGG 27110  
 Qy 360 TTCTGTCAACCCGTCGG 419  
 Db 27109 TTCTGTCAACCCGG 27050  
 Qy 420 AGCCCGCTTCGTCGG 479  
 Db 27049 CGCCCTGACTCGG 26990  
 Qy 480 CAGCAGACACCCAGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 539  
 Db 26989 CAGCAGACCGCCAGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 26930  
 Qy 540 GCCGAGAGTGCGG 599  
 Db 26929 GCCGAGAGTGCGG 26870  
 Qy 600 TCAGGCTGCGG 659  
 Db 26869 CCAGGCTGAAGGGCAGGATG---GCGAGGAGCGCGCGCGCGCGCGCGCGCGCGG 26813  
 Qy 660 GCGCGCGATGCGAGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 719  
 Db 26812 GCGCGCGATGCGAGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 26753  
 Qy 720 GGACCCCGGATTCGAGTCCCTCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 779  
 Db 26752 GGACCCCGGATTCGAGTCCCTCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGG 26699  
 Qy 780 CAAGGAGCGCTCGGTTCCAGTT 803  
 Db 26692 CAAGGAGCGCTCGGTTCCAGTT 26669

RESULT 8  
 AC125993/c  
 LOCUS AC125993 237695 bp DNA linear HTG 10-MAY-2003  
 DEFINITION Rattus norvegicus clone CH230-74L11, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*  
 AC125993  
 VERSION AC125993.3 GI:30522839  
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_ENRICHED.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 237695)  
 Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alibrooks,S., Amin,A., Anguiano,D.,  
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chu,J.,  
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Frasca,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Georgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,  
 Hollins,B., Howell,S., Hulyk,S., Hume,J., Idler,D., Jackson,A.,  
 Johnson,B., Johnson,H., Johnson,B., Johnson,R., Jolivet,A.,  
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lorensuwa,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,  
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
 Nwokeneme,O., Olarunpungoon,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,  
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,  
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Shen,H.,  
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczyk,R., Woodson,H., Worley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstein,G., and Gibbs,R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 237695)  
 Worley,K.C.  
 Direct Submission  
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 237695)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 10, 2003 this sequence version replaced gi:23096451.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GGIZ  
 Center clone name: CH230-74L11  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 229641 bases at least Q40  
 Consensus quality: 231444 bases at least Q30







Thu Jan 6 08:43:22 2005

QY 780 CAAGAGCGCTGGTTCCAGTT 803  
 Db 236554 CAAGAGCGCTGGTTCCAGTT 236531

RESULT 10  
 AC108848/c  
 LOCUS  
 DEFINITION Mus musculus chromosome 5 clone RP23-300K5 map 5, WORKING DRAFT  
 SEQUENCE, 40 unordered pieces.  
 ACCESSION AC108848  
 VERSION AC108848.3 GI:45430126  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Birren,B., Nusbaum,C. and Lander,E.  
 1 (bases 1 to 220894)  
 Mus musculus chromosome 5, clone RP23-300K5  
 2 (bases 1 to 220894)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J.J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,  
 Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Lacombe,K., Lamazares,R.,  
 Landers,T., Lechoczy,J., Levine,R., Liu,G., Maclean,C.,  
 MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,  
 McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,  
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,  
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 220894)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J.J., Choepel,Y.,  
 Collamore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,  
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 13, 2004 this sequence version replaced gi:20336129.

NOTE: This is a 'working draft' sequence. It currently  
 consists of 40 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

750: contig of 750 bp in length  
 850: gap of 100 bp  
 2504: contig of 1654 bp in length  
 2604: gap of 100 bp  
 3324: contig of 720 bp in length  
 3424: gap of 100 bp  
 16586: contig of 13162 bp in length  
 16586: gap of 100 bp  
 27361: contig of 10675 bp in length  
 27361: gap of 100 bp  
 32363: contig of 5802 bp in length  
 33363: gap of 100 bp  
 50014: contig of 16651 bp in length  
 50114: gap of 100 bp  
 56261: contig of 6147 bp in length  
 56261: gap of 100 bp  
 57961: contig of 1600 bp in length  
 58061: gap of 100 bp  
 58816: contig of 755 bp in length  
 58816: gap of 100 bp  
 58917: gap of 100 bp  
 61014: contig of 2098 bp in length  
 61114: gap of 100 bp  
 61973: contig of 859 bp in length  
 61974: gap of 100 bp  
 62074: contig of 691 bp in length  
 62764: gap of 100 bp  
 62765: gap of 100 bp  
 62865: contig of 2093 bp in length  
 64958: gap of 100 bp  
 71102: contig of 6045 bp in length  
 71202: gap of 100 bp  
 73181: contig of 1979 bp in length  
 73281: gap of 100 bp  
 76204: contig of 2923 bp in length  
 76304: gap of 100 bp  
 78361: contig of 2057 bp in length  
 78461: gap of 100 bp  
 82083: contig of 3622 bp in length  
 82183: gap of 100 bp  
 83917: contig of 1734 bp in length  
 84017: gap of 100 bp  
 85816: contig of 1759 bp in length  
 85917: gap of 100 bp  
 86672: contig of 756 bp in length  
 86672: gap of 100 bp

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L21014  
 Center clone name: 300 K 5  
 Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 200069 bases at least Q40  
 Consensus quality: 206302 bases at least Q30  
 Consensus quality: 208218 bases at least Q20  
 Insert size: 209348; sum-of-contigs  
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs

\* 86773 90021: contig of 3249 bp in length  
\* 90022 90121: gap of 100 bp  
\* 90122 92259: contig of 2138 bp in length  
\* 92260 92359: gap of 100 bp  
\* 92360 93758: contig of 1399 bp in length  
\* 93759 93858: gap of 100 bp  
\* 93859 110546: contig of 16688 bp in length  
\* 110547 110646: gap of 100 bp  
\* 110647 115943: contig of 5297 bp in length  
\* 115944 116043: gap of 100 bp  
\* 116044 130880: contig of 14837 bp in length  
\* 130881 130981: gap of 100 bp  
\* 130982 159619: contig of 28639 bp in length  
\* 159620 159719: gap of 100 bp  
\* 159720 163316: contig of 3597 bp in length  
\* 163317 163416: gap of 100 bp  
\* 163417 164845: contig of 1429 bp in length  
\* 164846 164945: gap of 100 bp  
\* 164946 165989: contig of 1044 bp in length  
\* 165990 166089: gap of 100 bp  
\* 166090 168069: contig of 1980 bp in length  
\* 168070 168169: gap of 100 bp  
\* 168170 169662: contig of 1493 bp in length  
\* 169663 169762: gap of 100 bp  
\* 169763 196645: contig of 26883 bp in length  
\* 196646 196745: gap of 100 bp  
\* 196746 214411: contig of 17666 bp in length  
\* 214412 214511: gap of 100 bp  
\* 214512 215537: contig of 1026 bp in length  
\* 215538 215637: gap of 100 bp  
\* 216833 216933: contig of 1196 bp in length  
\* 216934 218252: contig of 1319 bp in length  
\* 218253 218352: gap of 100 bp  
\* 218353 220894: contig of 2542 bp in length.

## FEATURES

source

1. .220894  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="5"  
/map="5"  
/clone="RP23-300K5"  
/clone\_lib="RPCI-23 Female Mouse BAC"  
1. .750  
/note="assembly\_fragment"  
851. .2504  
/note="assembly\_fragment"  
2605. .3324  
/note="assembly\_fragment"  
3425. .16596  
/note="assembly\_fragment"  
16687. .27361  
/note="assembly\_fragment"  
27462. .33263  
/note="assembly\_fragment"  
33364. .50014  
/note="assembly\_fragment"  
50115. .56261  
/note="assembly\_fragment"  
56362. .57961  
/note="assembly\_fragment"

Query Match 45.3%; Score 578.8; DB 2; Length 220894;

Best Local Similarity 83.9%; Pred. No. 7e-110; Matches 737; Conservative 0; Mismatches 46; Indels 95; Gaps 5;

Qy 3 GCGGGCGAGGCGGGGACGACCCATGTTCCCGGCGAGCAGTTCACACCCCTGCCCGCA 62

Db 27759 GCGGGCGAGGCGGGAACGACCCATGTTCCCGGCGAGCAGTTCACACCCCTGCCCGCA 27700

Qy 63 TCCTTATCCGCGGCCACCAAGCGGGGATGGCTGGAGGTCGAGGCCAGGGGCTGCCG 122

Db 27699 TCCTTATCCGCGGCCACCAAGCGGGGATGGCTGGAGGTCGAGGCCAGGGGCTGCCG 27640  
Qy 123 ACCGCGCCCCCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA 182  
Db 27639 ACCGCGCCCCCTCTCTCCCGGCTACAGACAGCTCATGGCCGCGGAGTACGTGCA 27580  
Qy 183 CAGCACCCAGCGGGCAGAGCTCATGGCCCTCTGTTCGCGGATGGGTCCCGGTCGGTCAG 242  
Db 27579 CCGCACCCAGCGGGCAGAGCTCATGGCCCTCTGTTCGCGGATGGGTCCCGGTCGGTCAG 27520  
Qy 243 CAGCCGTGACCTCGGTGCGAGTGAACCCCGCGCGCAGCCCTCGGTGCGAGTGTTC--- 299  
Db 27519 CAGCCGTGACCTCGGTGCGAGTGAACCCCGCGCGCAGCCCTCGGTGCGAGTGTTC--- 27460  
Qy 300 ----- 299  
Db 27459 NNN 27400  
Qy 300 -----ACTCGGCGCGCACGCTGCAGCCCTGCAGGGTGCAGGCCAGCCGCC 345  
Db 27399 NNN 27340  
Qy 346 GACGCCGATCGGGTCTCTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGCA 405  
Db 27339 GACGCCGCGTTCGGTTCCTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGCA 27280  
Qy 406 TCCTGGCAGACCGTAGCCCGTCTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCTCTG 465  
Db 27279 TCCTGGCAGACCGTTCGGTTCCTGTCCGTGCA-CTTCTGTGGCT---CTCTCTCTG 27224  
Qy 466 GAGGTTGCGGAGGAGCAGCAGACCCACGAAAGGAGAGGGGAGCCCGGATCTCTCGGG 525  
Db 27223 GAGGTTGCGGG-GGCGAGCAGACCCAGGAAGGAGAGGGGAGCCCGGATCTCTCGGAG 27165  
Qy 526 ACCCGGGAACCGGAGCGAGAGGTGGCCCGGAGAAAGCGGTCCCGCAGCGGGAAGC 585  
Db 27164 ACCCGGGAACCGGAGCGGAGAGGTGGCCGTGAGGAAGCGGTCCCGCAGCGGGAAGC 27105  
Qy 586 GAGGAGGCGGATGTTCAAGGTCGAGGCGAGCGCGGTGGGAGCAGCAGCCACCGGAG 645  
Db 27104 GAGGAGGCGGATGTTCAAGGTCGAGGCGAGCGCGGTGGGAGCAGCAGCCACCGGAG 27045  
Qy 646 GACCGGAACAGTGTGGCGGAGTCTGAGGCTTGGAGCGGAGGAGCCATGCTCTGCC 705  
Db 27044 GACCGGAACAGTGTGGTGTGATGATGAGCTTCTGAGCTTGGAGTGTGCTCTCTCT 26985  
Qy 706 GCAGAGTGGCTCAGGACCCCGGTGATTCGAGTGCCTTCGAGACCGGCTCCCGGCAA 765  
Db 26984 GCAGAGTGGCTCAGGACCCCGGTGATTCGAGTGCCTTCGAGACCGGCTCCCGGCAA 26938  
Qy 766 AGCAGGAGCAGGACAAAGGAGCGCTGCTGTTTCCAGTT 803  
Db 26937 AGCAGGAGCAGGACAAAGGAGCTCTCTGCTTTCAGGT 26900  
RESULT 11  
LOCUS AV191416 1275 bp mRNA linear PRI 12-MAR-2003  
DEFINITION Homo sapiens zygote arrest 1 (Zarl) mRNA, complete cds.  
ACCESSION AV191416  
VERSION AV191416.1 GI:27808691  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1275)  
AUTHORS Wu, X., Viveiros, M.M., Eppig, J.J., Bai, Y., Fitzpatrick, S.L. and Matzuk, M.M.  
TITLE Zygote arrest 1 (Zarl) is a novel maternal-effect gene critical for the oocyte-to-embryo transition  
JOURNAL Nat. Genet. 33 (2), 187-191 (2003)  
MEDLINE 22447938







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 14:32:42 ; Search time 679 Seconds  
(without alignments)  
9872.622 Million cell updates/sec

Title: US-09-830-810A-1

Perfect score: 1277

Sequence: 1 aaggcggcgagcgcgaggga.....acaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1265	99.1	1276	3 AAD00294	Aad00294 Mouse ooc
2	1245.4	97.5	1260	10 ADJ63191	Adj63191 Mouse zyg
3	1234.2	96.6	1258	8 ABZ24589	Abz24589 Mouse ova
4	1234.2	96.6	1258	10 ADJ63164	Adj63164 Mouse zyg
5	962.2	75.3	1280	10 ADJ63200	Adj63200 Rat zygote
6	799.4	62.6	814	10 ADJ63208	Adj63208 Zygote ar
7	799.4	62.6	6873	8 ABZ24590	Abz24590 Mouse ova
8	799.4	62.6	6873	10 ADJ63203	Adj63203 Mouse zyg
9	799.4	62.6	6873	10 ADJ63174	Adj63174 Mouse zyg
10	753.8	59.0	4090	8 ABZ24591	Abz24591 Mouse ova
11	753.8	59.0	4090	10 ADJ63175	Adj63175 Mouse zyg
12	628	49.2	3959	10 ADJ63201	Adj63201 Rat zygote
13	352	27.6	1275	10 ADJ63193	Adj63193 Human zyg
14	273.6	21.4	305	10 ADJ63211	Adj63211 Zygote ar
15	205.8	16.1	1052	10 ADJ63194	Adj63194 Frog zyg
16	194.8	15.3	963	10 ADJ63198	Adj63198 Torafugu
17	194.8	15.2	1192	10 ADJ63196	Adj63196 Zebrafish
18	134.6	10.5	2075	8 ABZ24592	Abz24592 Human ova
19	134.6	10.5	6002	10 ADJ63204	Adj63204 Human zyg
20	134.6	10.5	7405	10 ADJ63176	Adj63176 Human zyg
21	98.8	7.7	123	10 ADJ63209	Adj63209 Zygote ar

22	72	5.6	105	10	ADJ63210	Adj63210 Zygote ar
23	58.8	4.6	114955	2	AAX53491	Aax53491 Human ade
24	57.8	4.5	2000	8	ADA71938	Ada71938 Rice gene
25	54.2	4.2	1032	8	ADA65887	Ada65887 Rice gene
26	54.2	4.2	1500	12	ADI42217	Adi42217 Plant tra
27	54.2	4.2	1500	12	ADO02720	Ado02720 Soybean o
28	54.2	4.2	4667	12	ADQ22939	Adq22939 Human sof
c 29 52 4.1 114955 2 AAX53491						
c 30 51.6 4.0 987 8 ABZ37562						
c 31 51.6 4.0 59816 8 ABZ37516						
c 32 51.6 4.0 59816 8 ABZ37515						
c 33 50.2 3.9 1673 12 ADI42224						
c 34 50.2 3.9 1673 12 ADO02727						
c 35 49.6 3.9 400 2 AAV44436						
c 36 49.6 3.9 400 2 AAV64545						
c 37 49.6 3.9 400 2 AAZ19134						
c 38 49.6 3.9 400 2 AAZ19346						
c 39 49.6 3.9 14800 6 ABL66291						
c 40 49.6 3.9 14800 12 ADP13447						
c 41 49.6 3.9 14835 6 AAS94858						
c 42 49 3.8 5235 12 ADM31027						
c 43 49 3.8 6792 12 ADM31025						
c 44 49 3.8 7175 2 AAQ84658						
c 45 49 3.8 7175 2 AAV42686						

#### ALIGNMENTS

RESULT 1  
AAD00294  
ID AAD00294 standard; cDNA; 1276 BP.  
AC AAD00294;  
XX  
DT 05-SEP-2000 (first entry)  
DE Mouse oocyte-specific O1-180 cDNA clone.  
XX  
KW Oocyte-specific; ovary; O1-180; mouse; gynaecological; treatment; screen;  
KW cell proliferative disorder; cell degenerative disorder; contraceptive;  
KW modulator; signalling pathway; human infertility; cancer; ovulation; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 28..1113  
FT /\*tag= a  
FT /product= "Mouse oocyte-specific protein, O1-180"  
XX  
PN WO200024755-A1.  
XX  
PD 04-MAY-2000.  
XX  
PF 28-OCT-1999; 99WO-US025209.  
XX  
PR 28-OCT-1998; 98US-0106020P.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Matzuk MM, Wang P;  
XX  
DR WPI; 2000-350684/30.  
XX  
DR P-PSDB; AAY70948.  
XX  
PT O1-180, O1-184 and O1-236 polypeptides and nucleic acids encoding them,  
PT useful for evaluating potential contraceptives to block ovulation in a  
PT reversible manner.  
XX  
PS Claim 2; Fig 1; 54pp; English.  
XX  
CC The present sequence is the cDNA encoding the mouse oocyte-specific  
CC protein O1-180, expressed in the oocytes of primary (one-layer) preantral







XX The present sequence is that of murine ovary-specific O1-180 cDNA  
 CC obtained from a mouse ovary cDNA library. O1-180 clones were initially  
 CC identified in a subtractive hybridisation screen using ovaries from Gdf9  
 CC knockout and wild-type mice. The O1-180 gene (see AB24590) on chromosome  
 CC 5 is also provided. Loss of O1-180 results in female infertility and  
 CC subfertility. The invention provides ovary-specific and oocyte-specific  
 CC murine and human O1-180, O1-184 and O1-236 polynucleotides and  
 CC polypeptides. These genes and their protein products appear to relate to  
 CC various cell proliferative or degenerative disorders, especially those  
 CC involving ovarian tumours, such as germ line tumours and granulosa cell  
 CC tumours, or infertility, such as premature ovarian failure. The invention  
 CC provides a method for detection of a cell proliferative or degenerative  
 CC disorder of the ovary, which is associated with the expression of O1-180,  
 CC O1-184 or O1-236. It also provides a method for treating such disorders  
 CC by using an agent which suppresses or enhances the respective activities  
 CC of O1-180, O1-184 or O1-236, and a method of screening for compounds that  
 CC interact and/or modulate the expression or activity of the ovary-specific  
 CC genes. These compounds are possible contraceptive or fertility enhancing  
 CC agents. The modulator is preferably a polypeptide, small molecule or  
 CC polynucleotide sequence  
 XX  
 SQ Sequence 1258 BP; 261 A; 392 C; 392 G; 213 T; 0 U; 0 Other;  
 Query Match 96.6%; Score 1234.2; DB 8; Length 1258;  
 Best Local Similarity 99.3%; Pred. No. 2.4e-283;  
 Matches 1250; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
 3 GGGGGGCGAGCGGGGAGCGACCCATGTTCCGGGAGCAGCTTCCACCCCTGCCGCA 62  
 1 GGGGGGCGAGCGGGGAGCGACCCATGTTCCGGGAGCAGCTTCCACCCCTGCCGCA 60  
 63 TCCTTATCCGAGCGGCCACCAAGCCGGGATGCTGGAGTTTCGGAGCCAGGGGTGCGG 122  
 61 TCCTTATCCGAGCGGCCACCAAGCCGGGATGCTGGAGTTTCGGAGCCAGGGGTGCGG 120  
 123 ACCGGCGCCCTCTCTCTCCCGCTACAGACAGCTCAGTCCGCGCGGAGTACGTGCA 182  
 121 ACCGGCGCCCTCTCTCTCCCGCTACAGACAGCTCAGTCCGCGCGGAGTACGTGCA 180  
 183 CAGCCACGCGGCGACAGCTCATGCGCTCTGCTGCGGATGGGTCCCGGTGCGTCA 242  
 181 CCGCCACGCGGCGACAGCTCATGCGCTCTGCTGCGGATGGGTCCCGGTGCGTCA 240  
 243 CAGCCGTGACGCTGCGGTGAGGTGAACCCGCGCGGAGCCCTCGGTGAGTGTCACT 302  
 241 CAGCCGTGACGCTGCGGTGAGGTGAACCCGCGCGGAGCCCTCGGTGAGTGTCACT 300  
 303 CGGGCGCGGCGACGCTGAGCTGCGGCTCTGCTGCGGATGGGTCCCGGTGCGTCA 362  
 301 CGGGCGCGGCGACGCTGAGCTGCGGCTCTGCTGCGGATGGGTCCCGGTGCGTCA 360  
 363 CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 422  
 361 CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGATCTCTGGCAGACCGTAGC 420  
 423 CCGGTCTCTGCTGAGCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482  
 421 CCGGTCTCTGCTGAGCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
 483 GCAGACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542  
 481 GCAGACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 543 GAGAGAGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602  
 541 GAGAGAGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 603 GGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 662  
 601 GGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
 663 GGCGATGAGTGTGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 722

661 GGCGATGAGTGTGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 723 CCCCGGTGATTCGGATGCCCTTCGACACAGAGGCTCCCGGCAAGACCGAGGAGGACAA 782  
 721 CCCCGGTGATTCGGATGCCCTTCGACACAGAGGCTCCCGGCAAGACCGAGGAGGACAA 780  
 783 GGAGCGCTGCGGTTTCAGTTCCTTAGAGCAGAGTAGTACGCTACTATCTGCAAGGACTG 842  
 781 GGAGCGCTGCGGTTTCAGTTCCTTAGAGCAGAGTAGTACGCTACTATCTGCAAGGACTG 840  
 843 CAAAATCCGCTGGAGAGCGCCCTATGTGGTGTGTGCGAGGCGACAGTAAGGTGTACT 902  
 841 CAAAATCCGCTGGAGAGCGCCCTATGTGGTGTGTGCGAGGCGACAGTAAGGTGTACT 899  
 903 TCAACAGTTCGCGAGTGTGAGAAATCCTTACAACTTACAGAGTGGAGGACATCA 962  
 900 TCAACAGTTCGCGAGTGTGAGAAATCCTTACAACTTACAGAGTGGAGGACATCA 959  
 963 CTTGTCAAGTGTGTAAGAACTAGATGTGCTGCCAGTCAAGTTCGCGAGTGGACC 1022  
 960 CTTGTCAAGTGTGTAAGAACTAGATGTGCTGCCAGTCAAGTTCGCGAGTGGACC 1019  
 1023 CTAAACGCGCCCATCGCAAGACTTGTGGGAGATGCAAGGACAAACGCTGTCTCTGCG 1082  
 1020 CTAAACGCGCCCATCGCAAGACTTGTGGGAGATGCAAGGACAAACGCTGTCTCTGCG 1079  
 1083 ACAGCAGCTTCAGCTTCAATATACATCTTTAGTGAGAGTGCAGAAACGCTTCTGCTAGATG 1142  
 1080 ACAGCAGCTTCAGCTTCAATATACATCTTTAGTGAGAGTGCAGAAACGCTTCTGCTAGATG 1139  
 1143 GGGCTAATGGAATGGACAAGTGAAGTCTTCTCCCTCTTCACTTCTCCCTTCCAAATTC 1202  
 1140 GGGCTAATGGAATGGACAAGTGAAGTCTTCTCCCTCTTCACTTCTCCCTTCCAAATTC 1199  
 1203 TTCAATGACAGACAGTGTTCCTTGGATATAAGCTGTGTAATAAGGATTTGCAACAA 1261  
 1200 TTCAATGACAGACAGTGTTCCTTGGATATAAGCTGTGTAATAAGGATTTGCAACAA 1258

RESULT 4

ADJ63164  
 ID ADJ63164 standard; DNA; 1258 BP.

XX AC ADJ63164;

XX DT 06-MAY-2004 (first entry)

XX DE Mouse zygote arrest 1 (Zarl) protein coding sequence #1.

XX KW ovary-specific protein; zygote arrest 1; Zarl; nucleoplasmin; Npm2;  
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;  
 gene; db; mouse; murine.

XX OS Mus musculus.

XX PN WO2003091400-A2.

XX PD 06-NOV-2003.

XX PF 23-APR-2003; 2003WO-US012720.

XX PR 26-APR-2002; 2002WO-US013245.

XX PR 17-SEP-2002; 2002US-0411262P.

XX PR 17-DEC-2002; 2002US-0434165P.

XX PR 13-JAN-2003; 2003US-0439781P.

XX PR 23-JAN-2003; 2003US-0442164P.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 (AMHP ) WYETH.

XX PA Matzuk MM, Wang P, Wu X, Bai Y;

XX PI Matzuk MM, Wang P, Wu X, Bai Y;

DR WPI; 2003-865579/80.  
XX P-PSDB; ADJ63165.  
PT New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful  
PT in inhibiting or enhancing early embryogenesis or treating cell  
PT proliferative or degenerative disorders which are mediated by OI-180, OI-  
XX 184 or OI-236 proteins.  
PS Disclosure; SEQ ID NO 1; 219pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of ovary-  
CC specific (also known as zygote arrest 1 - Zar1) proteins and  
CC nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the  
CC invention are useful for inhibiting or enhancing early embryogenesis by  
CC disturbing the maternal genome. The DNA and protein sequences are also  
CC useful for screening mutations in components of those signalling pathways  
CC that are associated with some forms of human infertility or  
CC gynaecological cancers. The present DNA sequence encodes a Zar1 protein.  
XX  
SQ Sequence 1258 BP; 261 A; 392 C; 392 G; 213 T; 0 U; 0 Other;  
Query Match 96.6%; Score 1234.2; DB 10; Length 1258;  
Best Local Similarity 99.3%; Pred. No. 2.4e-283;  
Matches 1250; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
OY 3 GCGGGCGAGGCGCGGACGACACCATGTTCCCGGAGACAGCTTCCACCCCTGCCCCGCA 62  
DB 1 GCGGGCGAGGCGCGGAGCGACCATGTTCCCGGAGACAGCTTCCACCCCTGCCCCGCA 60  
OY 63 TCCTTATCCGAGGCGACAAAGCGGGGATGGCTGGAGGTCGAGGCGGCGGCGG 122  
DB 61 TCCTTATCCGAGGCGACAAAGCGGGGATGGCTGGAGGTCGAGGCGGCGGCGG 120  
OY 123 ACCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTCGA 182  
DB 121 ACCGCGCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTCGA 180  
OY 183 CAGCCACCGAGGCGACAGCTCATGCGCGGATGGGTCCCGGTCGGTCAG 242  
DB 181 CCGCCACCGAGGCGACAGCTCATGCGCGGATGGGTCCCGGTCGGTCAG 240  
OY 243 CAGCGGTGACGTGCGGTGCGAGGTGAACCCGCGCGCGAGCCTCGGTGCGAGTGTCACT 302  
DB 241 CAGCGGTGACGTGCGGTGCGAGGTGAACCCGCGCGCGAGCCTCGGTGCGAGTGTCACT 300  
OY 303 CCGGCGCGCAGCTGACGCTGCGAGGTGCGAGGTCGCGAGCCCGGATCGGCTTC 362  
DB 301 CCGGCGCGCAGCTGCGAGGTGCGAGGTCGCGAGGTCGCGAGCCCGGATCGGCTTC 360  
OY 363 CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGGATCCCTGGCAGACCGTAGC 422  
DB 361 CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGGATCCCTGGCAGACCGTAGC 420  
OY 423 CCGGTTCTGTCGCTGACCTTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482  
DB 421 CCGGTTCTGTCGCTGACCTTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
OY 483 GCAGACCCACGAGGAGGAGGAGCGCGGATCCCTCGGGACCCCGGACCCGAGGCC 542  
DB 481 GCAGACCCACGAGGAGGAGGAGCGCGGATCCCTCGGGACCCCGGACCCGAGGCC 540  
OY 543 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602  
DB 541 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
OY 603 GGCTGACGAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662  
DB 601 GGCTGACGAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
OY 663 GGCGATGAGTGTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722  
DB 661 GGCGATGAGTGTGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

723 CCCGGTGATTCGGATGCCCTCGAGACCGAGCCTCCCGCAAAAGCAGCGAGCAGGACAA 782  
DB CCCGGTGATTCGGATGCCCTCGAGACCGAGCCTCCCGCAAAAGCAGCGAGCAGGACAA 780  
OY 783 GGAGCGCTGGTTTCCAGTTCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTG 842  
DB 781 GGAGCGCTGGTTTCCAGTTCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGACTG 840  
OY 843 CAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902  
DB 841 CAAAATCCGGTGGGAGAGCGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899  
OY 903 TCAAAACAGTTCTCGCGGAGTGTGTGAGAAATCCTTACAACCCCTTACAGAGTGGAGGACATCA 962  
DB 900 TCAAAACAGTTCTCGCGGAGTGTGTGAGAAATCCTTACAACCCCTTACAGAGTGGAGGACATCA 959  
OY 963 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCCAGTGCAGATTTTCGCCACGTTGGACC 1022  
DB 960 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCCAGTGCAGATTTTCGCCACGTTGGACC 1019  
OY 1023 CTAAACGCCCCATCGGCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082  
DB 1020 CTAAACGCCCCATCGGCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079  
OY 1083 ACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAGAGTTCGAAACCGTTTCTGCTAGATG 1142  
DB 1080 ACAGCACCTTCAGCTTCAAATACATCAATTTAGTGAGAGTTCGAAACCGTTTCTGCTAGATG 1139  
OY 1143 GGGCTAATGGAATGGACAAAGTGTGCTTCCCTCTTTCACCTCTTCCCTTTCCAAATTC 1202  
DB 1140 GGGCTAATGGAATGGACAAAGTGTGCTTCCCTCTTTCACCTCTTCCCTTTCCAAATTC 1199  
OY 1203 TTCTATGACAGACAGT 1261  
DB 1200 TTCTATGACAGACAGT 1258

RESULT 5  
ADJ63200  
ID ADJ63200 standard; DNA; 1280 BP.  
XX  
AC ADJ63200;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Rat zygote arrest 1 (Zar1) DNA sequence.  
XX  
KW ovary-specific protein; zygote arrest 1; Zar1; nucleoplasmin; Npm2;  
KW embryogenesis; mutation screening; infertility; gynaecological cancer;  
KW ds; rat.  
XX  
OS Rattus norvegicus.  
XX  
PN WO2003091400-A2.  
XX  
PD 06-NOV-2003.  
XX  
PF 23-APR-2003; 2003WO-US012720.  
XX  
PR 26-APR-2002; 2002WO-US013245.  
PR 17-SEP-2002; 2002US-0411262P.  
PR 17-DEC-2002; 2002US-0434165P.  
PR 13-JAN-2003; 2003US-0439781P.  
PR 23-JAN-2003; 2003US-0442164P.  
XX  
(BAYU ) BAYLOR COLLEGE MEDICINE.  
PA (AMHP ) WYETH.  
XX  
PI Matzuk MM, Wang P, Wu X, Bai Y;  
XX  
DR WPI; 2003-865579/80.  
XX  
PT New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful

PT in inhibiting or enhancing early embryogenesis or treating cell  
 PT proliferative or degenerative disorders which are mediated by O1-180, O1-  
 PT 184 or O1-236 proteins.  
 XX

Claim 1; SEQ ID NO 37; 219pp; English.

XX The invention comprises the amino acid and coding sequences of ovary-  
 CC specific (also known as zygote arrest 1 - Zar1) proteins and  
 CC nucleoplamin (Npm2) proteins. The DNA and protein sequences of the  
 CC invention are useful for inhibiting or enhancing early embryogenesis by  
 CC disturbing the maternal genome. The DNA and protein sequences are also  
 CC useful for screening mutations in components of those signalling pathways  
 CC that are associated with some forms of human infertility or  
 CC gynaecological cancers. The present nucleic acid represents a Zar1 DNA  
 CC sequence.  
 XX

Sequence 1280 BP; 264 A; 408 C; 396 G; 212 T; 0 U; 0 Other;

Query Match 75.3%; Score 962.2; DB 10; Length 1280;  
 Best Local Similarity 87.7%; Pred. No. 1.1e-218;  
 Matches 1110; Conservative 0; Mismatches 143; Indels 12; Gaps 5;

QY 3 GCGGGCGAGGCGGGAGCGACCCATGTTCCCGGAGGAGCTTCCACCCCTGCCGCA 62  
 DB 3 GCGGGCGAGGCGGGAGCGACCCATGTTCCCGGAGGAGCTTCCACCCCTGCCGCA 62  
 QY 63 TCCTTATCCGC---AGGCCACCAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGGCTG 119  
 DB 63 TCCTTATCCGC---AGGCCACCAAGCCGGGATGGCTGGAGGTTTCGAGCCAGGGGCTG 122  
 QY 120 CGGACCGCGCGCCCTCTCTTCCCGGCTACAGACAGCTCATGCGCGGAGTACGT 179  
 DB 123 CAGGCGCGAGCCCGCTCTCTTCCCGGCTACAGACAGCTCATGCGCGGAGTACGT 182  
 QY 180 CGACAGCCACGCGGCGACAGCTCATGCGCGCTCTCTTCCCGGAGTACGT 239  
 DB 183 TGACAGCTATCAGGAGCGAGCTCATGCGCGCTCTCTTCCCGGAGTACGT 242  
 QY 240 CAGCAGCGGTGACGCTGCGGTGAGTGAACCGCGCGCGGAGCTCTGGTTCAGTGTTC 299  
 DB 243 CAGCAGCGGCGAGCGCTGCGGTGAGTGAACCGCGCGCGGAGCTCTGGTTCAGTGTTC 302  
 QY 300 ACTCGGGCGCGCAGCGCTGCGGTGAGTGAACCGCGCGCGGAGCTCTGGTTCAGTGTTC 359  
 DB 303 GCTCGGGCGCGCAGCGCTGCGGTGAGTGAACCGCGCGCGGAGCTCTGGTTCAGTGTTC 362  
 QY 360 TTCCTGTCAACCCCGTGGCCACCGCGCGCGGAGTATCCCGGATCTTGGCAGACCTG 419  
 DB 363 TTCCTGTCAACCCCGTGGCCACCGCGCGCGGAGTATCCCGGATCTTGGCAGACCTG 422  
 QY 420 AGCGCGGTCTGCTCGGTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479  
 DB 423 CGCGCTGTACTCGCGCGTGTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482  
 QY 480 CAGCAGACACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539  
 DB 483 CAGCAGACACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542  
 QY 540 GCCGAGAGAGGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599  
 DB 543 GCCGAGAGAGGTGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602  
 QY 600 TCAGGCTGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659  
 DB 603 CCAGGCTGAAGGGCAGGATG---GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659  
 QY 660 GCGGCGCATGAGTCTGAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719  
 DB 660 GCGGCGCATGAGTCTGAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719  
 QY 720 GGACCCCGGTATTCGGATGCGCTTCGAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779  
 DB 720 GGACCCCGGTATTCGGATGCGCTTCGAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779

QY 780 CAAGGAGCGCGTGGTTCAGTTCCTTACGAGAGTACGGCTACTATCATCTGCAAGCA 839  
 DB 780 CAAGGAGCGCGTGGTTCAGTTCCTTACGAGAGTACGGCTACTATCATCTGCAAGCA 839  
 QY 840 CTGCAAAATCCGTTGGAGAGCGCTTATGTGGTGTGTGCGAGGGCACCAGTAAGGTGT 899  
 DB 840 CTGCAAAATCCGTTGGAGAGCGCTTATGTGGTGTGTGCGAGGGCACCAGTAAGGTGT 898  
 QY 900 ACTTCAACAGATTCTCCGAGTGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGCA 959  
 DB 899 ACTTCAACAGATTCTCCGAGTGTGTGAGAAATCTTACAAACCTTACAGAGTGGAGCA 958  
 QY 960 TCACCTGTCAAGTGTGTAAAGAACTAGATGTGCTGCCAGTCAAGTTCGCGACGTGG 1019  
 DB 959 TCACCTGTCAAGTGTGTAAAGAACTAGATGTGCTGCCAGTCAAGTTCGCGACGTGG 1018  
 QY 1020 ACCTAAACCGCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAACGCTGTCTCT 1079  
 DB 1019 ACCTAAACCGCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAACGCTGTCTCT 1078  
 QY 1080 GCGACAGCAGCTTCAAGTTCAAATATCATATTTAGTGGAGTCAAAACGTTTCTGCTA- 1138  
 DB 1079 GTGACAGCAGCTTCAAGTTCAAATATCATATTTAGTGGAGTCAAAACGTTTCTGCTA- 1138  
 QY 1139 ---GATGGGCTAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1195  
 DB 1139 ATGATGCGGCAAGTGGAGCTTCTCCACGCGCTCTCCCTCTCCCTCTCCCTCTCTCT 1198  
 QY 1196 CAAATCTTATGACAGCAGTGTGTCTGGATATTAAGCTGTGAATAAAGGTATTGC 1255  
 DB 1199 AATACTTCAAGAGGAGTGT-TACTTTAATAAAGCTGTGAATAAAGGTATTGC 1257  
 QY 1256 AAACA 1260  
 DB 1258 AAACA 1262

RESULT 6  
 ADJ63208  
 ID ADJ63208 standard; DNA; 814 BP.  
 XX  
 AC ADJ63208;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Zygote arrest 1 (Zar1) gene - exon 1.  
 XX  
 KW ovary-specific protein; zygote arrest 1; Zar1; nucleoplasmin; Npm2;  
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;  
 KW ds; exon 1.  
 OS Unidentified.  
 XX  
 PN WO2003091400-A2.  
 XX  
 PD 06-NOV-2003.  
 XX  
 XX 23-APR-2003; 2003WO-US012720.  
 XX  
 XX 26-APR-2002; 2002WO-US013245.  
 PR 17-SEP-2002; 2002US-0411262P.  
 PR 17-DEC-2002; 2002US-0434165P.  
 PR 13-JAN-2003; 2003US-0439781P.  
 PR 23-JAN-2003; 2003US-0442164P.  
 XX  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 PA (AMHP ) WYETH.  
 XX  
 XX Matzuk MM, Wang P, Wu X, Bai Y;  
 XX WPI; 2003-865579/80.  
 XX















XX WO2003091400-A2.  
 XX 06-NOV-2003.  
 XX 23-APR-2003; 2003WO-US012720.  
 XX 26-APR-2002; 2002WO-US013245.  
 XX 17-SEP-2002; 2002US-0411262P.  
 XX 17-DEC-2002; 2002US-0434165P.  
 XX 13-JAN-2003; 2003US-0439781P.  
 XX 23-JAN-2003; 2003US-0442164P.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX (AMHP ) WYETH.  
 XX Matzuk MM, Wang P, Wu X, Bai Y;  
 XX WPI; 2003-865579/80.  
 XX P-PSDB; ADJ63202.  
 XX New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful  
 XX in inhibiting or enhancing early embryogenesis or treating cell  
 XX proliferative or degenerative disorders which are mediated by OI-180, OI-  
 XX 184 or OI-236 proteins.  
 XX Claim 1; SEQ ID NO 38; 219pp; English.  
 XX The invention comprises the amino acid and coding sequences of ovary-  
 XX specific (also known as zygote arrest 1 - Zar1) proteins and  
 XX nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the  
 XX invention are useful for inhibiting or enhancing early embryogenesis by  
 XX disturbing the maternal genome. The DNA and protein sequences are also  
 XX useful for screening mutations in components of those signalling pathways  
 XX that are associated with some forms of human infertility or  
 XX gynaecological cancers. The present DNA sequence encodes a Zar1 protein.  
 XX Sequence 3959 BP; 936 A; 1014 C; 1043 G; 966 T; 0 U; 0 Other;  
 XX  
 XX Query Match 49.2%; Score 628; DB 10; Length 3959;  
 XX Best Local Similarity 88.1%; Pred. No. 4.8e-139;  
 XX Matches 708; Conservative 0; Mismatches 90; Indels 6; Gaps 2;  
 QY 3 GCGGGCGGGCGGGAGCGCACCATTGTCGGGCGAGCAGCTTCCACCCCTGCCGCA 62  
 DB 25 GCGGGCGGGCGGGAGCGCACCATTGTCGGGCGAGCAGCTTCCACCCCTGCCGCA 84  
 QY 63 TCCTTATCCGC---AGGCCACCAAGCCGGGATGCTCGAGTTCCGAGCCAGGGGTG 119  
 DB 85 TCCTTATCCGC---AGGCCACCAAGCCGGGATGCTCGAGTTCCGAGCCAGGGGTG 144  
 QY 120 CCGACCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 179  
 DB 145 CAGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 204  
 QY 180 CGACAGCCACAGCGGCGCAGCTCATGCGCCCTGCTGTCGGATGGGTCCCGGTCGGT 239  
 DB 205 TGACAGCTATCAGGAGCGGAGCTCATGCGCCCTGCTGTCGGATGGGTCCCGGTCGGT 264  
 QY 240 CAGCAGCGGTGACGCTCGGTGACGTGAACCCGCGCGGCGAGCCCTCGGTGACGTGTC 299  
 DB 265 CAGCAGCGGTGACGCTCGGTGACGTGAACCCGCGCGGCGAGCCCTCGGTGACGTGTC 324  
 QY 300 ACTCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 359  
 DB 325 GCTCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 384  
 QY 360 TTCCTGTCAACCCCGTGGCCACGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 419  
 DB 385 TTCTGTCAACCCCGTGGCCACGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 444  
 QY 420 AGCCCGGTTCGTGCGGTGACCTTCTGTGGCCCTCTCTCTCTCTCTCTCTCTCTCT 479

Db 445 CGCCTGTACTCGCCCGTGACCTTCGCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 504  
 QY 480 CAGCAGACACCCACGAAAGGAGAGGGAGCCCGGATCTCTCGGGGACCCCGGAAACCGGA 539  
 Db 505 CAGCAGACACCCACGAAAGGAGAGGGAGCCCGGATCTCTCGGGGACCCCGGAAACCGGA 564  
 QY 540 GCCAGAGAGTGGCCCGGAGGAAAGCGTCCCGACGCGGAGAGCGAGGAGGAGGCGATGT 599  
 Db 565 GCCGGGAGAGTGGCAGTGAAGAGCAGTCCCGACGCGGAGAGCGAGGAGGCGGACGT 624  
 QY 600 TCAGGCTGCAGGCGAGCCCGGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGT 659  
 Db 625 CCAGGCTGAAGGGCAGGATG---GCCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGT 681  
 QY 660 GCGCGGATGCACTCTGAGCCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 719  
 Db 682 GCGCGGATGCACTCTGAGCCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 741  
 QY 720 GGACCCCGGTGATTTCGGATGCCCTCGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 779  
 Db 742 GGACCCCGGTGATTTCGGATGCCCTCGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 801  
 QY 780 CAAGGAGCGCTGCGTTCAGATT 803  
 Db 802 CAAGGAGCGCTGCGTTCAGATT 825

RESULT 13  
 ID ADJ63193  
 AC ADJ63193;  
 DT 06-MAY-2004 (first entry)  
 DE Human zygote arrest 1 (Zar1) DNA sequence #2.  
 KW ovary-specific protein; zygote arrest 1; Zar1; nucleoplasmin; Npm2;  
 KW embryogenesis; mutation screening; infertility; gynaecological cancer;  
 KW ds; human.  
 OS Homo sapiens.  
 PN WO2003091400-A2.  
 PD 06-NOV-2003.  
 XX 23-APR-2003; 2003WO-US012720.  
 XX 26-APR-2002; 2002WO-US013245.  
 XX 17-SEP-2002; 2002US-0411262P.  
 XX 17-DEC-2002; 2002US-0434165P.  
 XX 13-JAN-2003; 2003US-0439781P.  
 XX 23-JAN-2003; 2003US-0442164P.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX (AMHP ) WYETH.  
 XX Matzuk MM, Wang P, Wu X, Bai Y;  
 XX WPI; 2003-865579/80.  
 XX New isolated ovary-specific genes e.g. OI-180, OI-184 or OI-236, useful  
 XX in inhibiting or enhancing early embryogenesis or treating cell  
 XX proliferative or degenerative disorders which are mediated by OI-180, OI-  
 XX 184 or OI-236 proteins.  
 XX Claim 1; SEQ ID NO 30; 219pp; English.  
 XX The invention comprises the amino acid and coding sequences of ovary-  
 XX specific (also known as zygote arrest 1 - Zar1) proteins and  
 XX nucleoplasmin (Npm2) proteins. The DNA and protein sequences of the  
 XX invention are useful for inhibiting or enhancing early embryogenesis by



Qy	965	TGTCAAAGTTGTGTA	AAAGAACTAGAT	TGCTGCCAGTCAGAT	TTCGCCACGTGGAC	CCCT	1024
Db	10	TTTTCAGAGTTGT	TAAAGAACTAGAT	TGCTGCCAGTCAGAT	TTCGCCACGTGGAC	CCCT	69
Qy	1025	AAACGCCCCCAT	CGGCAAGACTTGT	TGGGAGATGCA	AGGACAAAACCCCTGT	CCTGCGAC	1084
Db	70	AAACGCCCCCAT	CGGCAAGACTTGT	TGGGAGATGCA	AGGACAAAACCTGT	CCTGCGAC	129
Qy	1085	AGCACCTTCAGCTT	CAAAATCATCAT	TTATGTGAGAGTC	GAAAAACGTTTTCTGCTAGAT	TGGG	1144
Db	130	AGCACCTTCAGCTT	CAAAATCATCAT	TTATGTGAGAGTC	GAAAAACGTTTTCTGCTAGAT	TGGG	189
Qy	1145	GCTAATGGAA	TGGACAGTCAGCTT	CTCCCTCTTCA	CCCTCTTCCCTTTCCAAAT	TCTT	1204
Db	190	GCTAATGGAA	TGGACAGTCAGCTT	CTCCCTCTTCC	CTTCCCAATTTCCAAAT	TCTT	249
Qy	1205	CATGACAGACAGT	TGTACTTGGATATA	AAAGCCTGTGAAT	AAAAAGGTATTCCAAACA		1260
Db	250	CATGACAGACAGT	TGTACTTGGATATA	AAAGCCTGTGAAT	AAAAAGGTATTCCAAACA		305

RESULT 15

ADJ63194

ID ADJ63194 standard; DNA; 1052 BP.

AC ADJ63194;

06-MAY-2004 (first entry)

XX  
DE  
Frog zygote arrest 1 (Zar1) protein coding sequence.

AA  
KW  
KW  
KW  
KW

AA  
OS Xenopus laevis.

AX WO2003091400-A2.  
PN

06-NOV-2003.

23-APR-2003: 2003WO-US012720-XX PF

XX  
PR 26-APR-2002. 2002WO-IIS013245

```
PR 17-SEP-2002; 2002US-0411262P
PR 17-DEC-2002; 2002US-0434165P
```

PR 13-JAN-2003; 2003US-0439781P

XX

РА (АМНР) WYETH.

PI Matzuk MM, Wang P, Wu X, Bai

XX  
DR WPI; 2003-865579/80.

DR F-FSDB; ADJ63193.  
XX

PT in inhibiting or enhancing early e

PT proliferative or degenerative disease  
DM 194 or 01-225 proteinase

XX  
pg  
[aim 1: 890 TD NO 21. 210mm. Enc]

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E

XX

specific (also known as zygote arr

CC invention are useful for inhibiting

CC useful for screening mutations in

CC gynaecological cancers. The present

XX	SQ	Sequence	1052 BP;	298 A;	228 C;	299 G;	227 T;	0 U;	0 Other;
		Query Match	16.1%;	Score	205.8;	DB	10;	Length	1052;
		Best Local Similarity	80.1%;	Pred. No.	7.7e-39;				
		Matches	254;	Conservative	0;	Mismatches	62;	Indels	1; Gaps 1;
Qy	796	TTCCAGTTCCTTAGACGAGAGTAGTCGGCTACTATCATCTGCAAGGACTGC	AAATCCGGTGG	855					
Db	594	TTCCAGTTCCTGAGCAGAGATACGGATATTACCTGTAAGGACTGC	ACATCCGGTGG	653					
Qy	856	GAGAGCGCCCTATGTGTGGTGTGTCAGGGCCACAGTAAGTGTGTACTTCAAACAGTTC	TCTG	915					
Db	654	GAGAGCGCCCTACGTGTGGTGTGTCAGGAAACCAATAAGGTG-TACTTCAAGCAGTTC	TCTG	712					
Qy	916	CCGAGTGTGTGAAATCCTTACAAACCTTACAGATGGAGGACATCACCTGTCAAAGTTG		975					
Db	713	CAGGACATGTGCAGAAATCCTTAAATCCCTACCGGTGGGAAGCATCATGTGTGAGAGCTG		772					
Qy	976	TAAAGAACTAGATGTGCTGCCAGTCAAGATTCGCCACGTGGACCTTAAACGCCGCCA		1035					
Db	773	CAGCAGACGAGATGCGGTGTCTGTCAAACCTGCGTCACGTTCAGCCCAAGAGGCCCA		832					
Qy	1036	TCGGCAAGACTTGTGTGGGAGATGCAAGGACAAACGCCCTGTCTCTGGCAGACGACCTTCAG		1095					
Db	833	CGCCAGGATCTGTGTGGGAGATGCAAGCAACGGCTCTCTGTGTGACAGCACTTTTAG		892					
Qy	1096	CTTCAAATCATCATTT	1112						
Db	893	CTTCAAGTATATCATTT	909						

Search completed: January 5, 2005, 16:34:49  
Job time : 698 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2005, 16:10:41 ; Search time 128 Seconds  
(without alignments)  
7091.230 Million cell updates/sec

Title: US-09-830-810A-1  
Perfect score: 1277  
Sequence: 1 aagcggcgagcgccgga.....acaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	4.2	5467	1 US-07-745-206A-12	Sequence 12, Appl
2	54	4.2	5467	2 US-08-311-363-12	Sequence 12, Appl
3	53.2	4.2	7218	1 US-08-232-463-14	Sequence 14, Appl
4	50.8	4.0	6232	3 US-08-456-200B-11	Sequence 11, Appl
5	49.6	3.9	400	3 US-09-056-556-179	Sequence 179, Appl
6	49.6	3.9	400	4 US-09-072-596-174	Sequence 174, Appl
7	49.6	3.9	400	4 US-09-072-967-179	Sequence 179, Appl
8	49	3.8	7175	1 US-08-455-543A-8	Sequence 8, Appl
9	49	3.8	7175	2 US-08-193-078B-8	Sequence 8, Appl
10	49	3.8	7175	2 US-08-223-305C-8	Sequence 8, Appl
11	49	3.8	7175	2 US-08-149-097D-8	Sequence 8, Appl
12	49	3.8	7175	3 US-08-949-386-8	Sequence 8, Appl
13	49	3.8	7175	3 US-08-450-582-8	Sequence 8, Appl
14	49	3.8	7175	3 US-08-984-709A-8	Sequence 8, Appl
15	49	3.8	7175	3 US-08-450-272-8	Sequence 8, Appl
16	49	3.8	7175	4 US-08-450-273-8	Sequence 8, Appl
17	49	3.8	7177	3 US-09-268-163-7	Sequence 7, Appl
18	49	3.8	7266	3 US-08-713-118-1	Sequence 1, Appl
19	49	3.8	7266	3 US-09-452-007-1	Sequence 1, Appl
20	49	3.8	7362	1 US-08-455-543A-7	Sequence 7, Appl
21	49	3.8	7362	2 US-08-193-078B-7	Sequence 7, Appl
22	49	3.8	7362	2 US-08-223-305C-7	Sequence 7, Appl
23	49	3.8	7362	3 US-08-149-097D-7	Sequence 7, Appl
24	49	3.8	7362	3 US-08-949-386-7	Sequence 7, Appl
25	49	3.8	7362	3 US-08-450-582-7	Sequence 7, Appl
26	49	3.8	7362	3 US-08-984-709A-7	Sequence 7, Appl
27	49	3.8	7362	3 US-08-450-272-7	Sequence 7, Appl

28	49	3.8	7362	4 US-08-450-273-7	Sequence 7, Appl
29	49	3.8	7364	3 US-09-268-163-5	Sequence 5, Appl
30	49	3.8	7376	3 US-09-268-163-3	Sequence 3, Appl
31	48	3.8	4403785	3 US-09-103-840A-2	Sequence 2, Appl
32	46.6	3.6	745	1 US-08-036-555B-163	Sequence 163, App
33	46.6	3.6	745	1 US-08-469-569-163	Sequence 163, App
34	46.6	3.6	745	1 US-08-249-322A-163	Sequence 163, App
35	46.6	3.6	745	1 US-08-469-526A-163	Sequence 163, App
36	46.6	3.6	745	2 US-08-734-591A-163	Sequence 163, App
37	46.6	3.6	745	2 US-08-469-660-163	Sequence 163, App
38	46.6	3.6	745	3 US-08-341-018-51	Sequence 51, Appl
39	46.6	3.6	745	3 US-08-470-335-163	Sequence 163, App
40	46.6	3.6	745	3 US-08-735-021-163	Sequence 163, App
41	46.6	3.6	745	3 US-08-734-664A-163	Sequence 163, App
42	46.6	3.6	745	3 US-08-470-339-163	Sequence 163, App
43	46.6	3.6	745	4 US-08-467-602-163	Sequence 163, App
44	46.6	3.6	745	4 US-08-411-295F-44	Sequence 44, Appl
45	46.6	3.6	745	5 PCT-US94-05083C-159	Sequence 159, App

## ALIGNMENTS

RESULT 1  
US-07-745-206A-12  
; Sequence 12, Application US/07745206A  
; Patent No. 5429921  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: McCue, Ann  
; APPLICANT: Feldman, Daniel  
; TITLE OF INVENTION: Human Calcium Channel Compositions and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fitch, Even, Tabin & Flannery  
; STREET: 135 S. LaSalle  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07745,206A  
; FILING DATE: 19910815  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feder, Scott B  
; REFERENCE/DOCKET NUMBER: 51504  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-372-7842  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5467 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390  
; LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585  
; LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744  
; LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010  
; LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)  
US-07-745-206A-12











```
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGGTGACAGTGAACCCGCGCGCGCGAGCTCTCGTGCACTGTTCACTCGGGCGC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGTGTCCCGGGAGGAGCGGCCCGCGCGCAC 2810
QY 310 CCGACGCTGACCTTGTGAGGTGCGGAGCGAGCCCGAGCCCGCATCGGTTCCTGTCAA 369
Db 2811 CGCA-GCCACACAAGGAGCGCGCGGGCC--CCCGAGGCGCGGAGCGAGCGCGCGCA 2867
QY 370 CCCCGTGGCGACCGCGCGCGGAGATCCCGCGCATCTGSCACACGTCAGCCCGTTC 429
Db 2868 GCGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCGCTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCACAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGGAGAAAGCGGTCTCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 609
Db 3048 GCGGAGAGCGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GGGCAGGCGCGGTGGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
Db 3108 CACGAGGCTGTGAGAGAGGAGCACCAGGAGGAGCGCGCGCGCGCGCGCGCGCGAT 3166

RESULT 9
US-08-193-078B-8
; Sequence 8, Application US/08193078B
; Patent No. 5846757
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
; STREET: 1660 UNION STREET
; CITY: SAN DIEGO
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,078B
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-53607
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
```

```
TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
US-08-193-078B-8

Query Match 3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0091;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGGTGACAGTGAACCCGCGCGCGCGAGCTCTCGTGCACTGTTCACTCGGGCGC 309
Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGTGTCCCGGGAGGAGCGGCCCGCGCGCAC 2810
QY 310 CCGACGCTGACCTTGTGAGGTGCGGAGCGAGCCCGAGCCCGCATCGGTTCCTGTCAA 369
Db 2811 CGCA-GCCACACAAGGAGCGCGCGGGCC--CCCGAGGCGCGGAGCGAGCGCGCGCGA 2867
QY 370 CCCCGTGGCGACCGCGCGCGGAGATCCCGCGCATCTGSCAGACCGTCAGCCCGTTC 429
Db 2868 GCGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2927
QY 430 TCGTCGCTGACCTTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
QY 490 CCCACAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2988 GCGCCAAAGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
QY 550 GTGGCGCGGAGAAAGCGGTCTCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 609
Db 3048 GCGGAGAGCGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GGGCAGGCGCGGTGGGAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 668
Db 3108 CACGAGGCTGTGAGAGAGGAGCACCAGGAGGAGCGCGCGCGCGCGCGCGCGCGAT 3166

RESULT 10
US-08-223-305C-8
; Sequence 8, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/223,305C
  FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/868,354
  FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/745,206
  FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/620,250
  FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/482,384
  FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/603,751
  FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: WO PCT/US89/01408
  FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/176,899
  FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
  NAME: Seidman, Stephanie L.
  REGISTRATION NUMBER: 33,779
  REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
  LENGTH: 7175 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
  NAME/KEY: CDS
  LOCATION: 144..6857
FEATURE:
  NAME/KEY: 5'UTR
  LOCATION: 1..143
FEATURE:
  NAME/KEY: 3'UTR
  LOCATION: 6855..7175
US-08-223-305C-8

Query Match      3.8%; Score 49; DB 2; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0091;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGTGCAGGTGAACCCGCGCGCGGAGCGCTCGGTGCAGTGTTCCTCGGGCCG 309
DB 2751 GAGGCCCCGAGGCGGAGGCGGGAGCGCCGCGTGCCTCGGAGGAGCGCGCCGCGCGAC 2810
QY 310 CGCAGCTGCAGCTGCAGGGTGCAGCGCCGAGCGCCCGGATCGGTTCTGTCTCA 369
DB 2811 CGCA-GCCACAGCAGAGGAGCGCGGGGCC--CCCGAGGCGCGGAGCGGCGGCCGA 2867
QY 370 CCCGTCGCCACGCGCGCGCGGAGATCCCGCGGATCCTGCGACAGCCGTAGCCCGGTC 429
DB 2868 GSCCCAGGCCCGAGGCGCGCGCGGCGCACACCGCGCGGCTCCCGGAGGAGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTTCCTCTCTGAGAGTTGCGGAGGAGGAGGACACA 489
DB 2928 GAGCGGGAGCCCCGACCGCGCGCGGACCGGACCGGATCCGAGGAGGAGTGCGCC 2987

QY 490 CCCACGAGGAGGAGGAGCGCGCATCTCTCGGGACCCCGGAACCGGAGCCGAGAGAG 549
DB 2988 GCGCCCAAGGCGGAGCGCGCGCGCGGCGCACCGCGCGGCCCCCGAGCGGCGCGGAG 3047
QY 550 GTGCGCGGAGGAAAGCGGTCTCCCGCAGCGCGGAAAGCGAGGAGGCGGCGATGTTTCAGGCTGCA 609
DB 3048 GCGGAGAGCGGGGAGGAGCGCGCGCGGCGGCGCACCGGGCCCGGCAACAAGCGCGAGCTGCT 3107
QY 610 GGGCAGGCGCGGTGGGAGCAGCAGCCACCACCGGAGGACCGGAAACAGTGTGGCGGCGAT 668
DB 3108 CACGAGGCTGTGGAGAGGAGACCA CGGAGAGGAGGCGCACCGGAGAGGAGGCTGAGAT 3166

RESULT 11
US-08-149-097D-8
; Sequence 8, Application US/08149097D
; Patent No. 5874236
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,097D
; FILING DATE: 05-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/06903
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,231
; FILING DATE: 13-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; APPLICATION NUMBER: US 07/176,899
```

Thu Jan 6 08:43:23 2005

FILED DATE: 04-APR-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-55038  
TELEPHONE: (619) 238-0999  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-149-097D-8

Query Match 3.8%; Score 49; DB 2; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0091;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GAGCGTGGGTGAGTGAACCGCGCCGCGAGCGTCCGTCAGTTCCTCGGGCGC 309  
DB 2751 GAGGCCCCGAAGCGGAGCGGCGGAGCGGCGGAGCGGCGGCGGCGGCGC 2810  
QY 310 CGCAGCTGCGAGTTCGAGGGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGC 369  
DB 2811 CGCA-GCCACAGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2867  
QY 370 CCCCGTGGCGACCTTCTGTGGGCTCTCTCTCACTGAGGTTGCGGAGCGGCGGCGC 429  
DB 2868 GCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2927  
QY 430 TCGTCCGTGACCTTCTGTGGGCTCTCTCTCACTGAGGTTGCGGAGCGGCGGCGC 489  
DB 2928 GAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2987  
QY 490 CCCACGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 549  
DB 2988 GCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 3047  
QY 550 GTGGCCCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 609  
DB 3048 GCGGAGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 3107  
QY 610 GGGCAGCGCGGCGGCGGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGC 668  
DB 3108 CACGAGGCTGTGGAGAGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGC 3166

RESULT 12  
US-08-949-386-8  
Sequence 8, Application US/08949386  
Patent No. 6090623  
GENERAL INFORMATION:  
APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: McCue, Ann  
APPLICANT: Gillespie, Allison  
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: US  
ZIP: 92101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/949,386  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,012  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 08/149,097  
FILING DATE: 5-NOV-1993  
PRIOR APPLICATION DATA: 08/105,536  
APPLICATION NUMBER: 11-AUG-1993  
FILING DATE: 11-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 519808  
TELEPHONE: (619) 238-0062  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7175 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 144..6857  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..143  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 6855..7175  
US-08-949-386-8

Query Match 3.8%; Score 49; DB 3; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0091;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;  
QY 250 GAGCGTGGGTGAGTGAACCGCGCCGCGAGCGTCCGTCAGTTCCTCGGGCGC 309  
DB 2751 GAGGCCCCGAAGCGGAGCGGCGGAGCGGCGGAGCGGCGGCGGCGGCGC 2810  
QY 310 CGCAGCTGCGAGTTCGAGGGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGC 369  
DB 2811 CGCA-GCCACAGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2867  
QY 370 CCCCGTGGCGACCTTCTGTGGGCTCTCTCTCACTGAGGTTGCGGAGCGGCGGCGC 429  
DB 2868 GCGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2927  
QY 430 TCGTCCGTGACCTTCTGTGGGCTCTCTCTCACTGAGGTTGCGGAGCGGCGGCGC 489  
DB 2928 GAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 2987  
QY 490 CCCACGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 549  
DB 2988 GCGGCGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 3047  
QY 550 GTGGCCCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGC 609



Db 3108 CACGAGGCTGTGGAGAGAGACACCGAGAGAGGAGCCACCGAGAGAGAGGCTGAGAT 3166

## RESULT 14

US-08-984-709A-8  
; Sequence 8, Application US/08984709A  
; Patent No. 6320032  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Mark E.  
; APPLICANT: Stauderman, Kenneth A.  
; APPLICANT: Harpold, Michael M.  
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS  
; CORRESPONDENCE ADDRESS:  
; NUMBER OF SEQUENCES: 52  
; ADDRESSEE: Heller Ehrman White & McAuliffe  
; STREET: 4250 Executive Square, Suite 700  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,709A  
; FILING DATE: 02-DEC-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seidman, Stephanie L.  
; REGISTRATION NUMBER: 33,779  
; REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 450-8400  
; TELEFAX: (619) 587-5360  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7175 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 144..6857  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: 1..143  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: 6855..7175  
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 3; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0091;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;  
QY 250 GACGTGCGGTGCAGGTGAACCGCGCCGACGCTCGGTGCAGTGTTCACCTCGGGCGC 309  
Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGTGTCCCGGAGGAGCGCGCGCGGCAC 2810  
QY 310 CGCAGCTCAGCCTGCAGGGTCCGAGCCAGCCCGACGCCCGATCGGGTTCCTGTCAA 369  
Db 2811 CGCA-GCCACAGCAAGAGGCGCGGGGCC--CCCGAGGCGCGGAGCGAGCGCGGCCA 2867  
QY 370 CCCCGTGGCAGCCCGCGCGGAGATCCCGCGATCCTGGCAGACCGTAGCCCGGTTTC 429  
Db 2868 GGCACAGGCG 2927  
QY 430 TCGTCCGTGACCTTCTGTGGCCTCTCTCTCTCACTGGAGGTTTGGCGGGAGGCGAGCA 489

Db 2928 GAGCGGAGCCCCGACGCCACCGCGCCACCGGACACGAGATCCGAGCAAGAGTGCGCC 2987  
QY 490 CCCACGAAGGAGAGGGAGCCCGGCATCTCTCGGGGACCCCGGGAACCGGAGCCGAGAGAG 549  
Db 2988 GCGCCAAAGCGGAGCGCGCGCGCGGCACCGCGCGCGCGCCCGAGCGGGCGCGCGGAG 3047  
QY 550 GTGGCCCGGAGAAAGCGGTCCCGACCGCGGAGCGAGAGGGGCGGATGTTTCAGGCTCA 609  
Db 3048 GCGGAGAGCGGGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3107  
QY 610 GGGCAGGCGCGGTGGGAGCAGCAGCCACCGGAGGAGCCGGAACAGTGTGTGGCGCGAT 668  
Db 3108 CACGAGGCTGTGGAGAGGAGACCCACCGAGAGGAGGCCACCGAGAGAGGAGGCTGAGAT 3166

## RESULT 15

US-08-450-272-8  
; Sequence 8, Application US/08450272  
; Patent No. 6387696  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: McCue, Ann  
; APPLICANT: Gillespie, Alison  
; APPLICANT: Feldman, Daniel  
; APPLICANT: Brenner, Robert  
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street  
; CITY: San Diego  
; STATE: California  
; COUNTRY: US  
; ZIP: 92101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,272  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/404,950  
; FILING DATE: 13-MAR-1995  
; APPLICATION NUMBER: 08/336,257  
; FILING DATE: 7-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/314,083  
; FILING DATE: 28-SEPT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/311,363  
; FILING DATE: 23-SEPT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/290,012  
; FILING DATE: 11-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/223,305  
; FILING DATE: 4-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/193,078  
; FILING DATE: 07-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/149,097

;; FILING DATE: 5-NOV-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/105,536  
;; FILING DATE: 11-AUG-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/914,231  
;; FILING DATE: 13-JULY-1992  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/868,354  
;; FILING DATE: 10-APR-1992  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US92/06903  
;; FILING DATE: 14-AUG-1992  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/745,206  
;; FILING DATE: 15-AUG-1991  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Seidman, Stephanie L.  
;; REGISTRATION NUMBER: 33,779  
;; REFERENCE/DOCKET NUMBER: 6362-519812  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 238-0999  
;; TELEFAX: (619) 238-0062  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7175 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 144..6857  
;; FEATURE:  
;; NAME/KEY: 5'UTR  
;; LOCATION: 1..143  
;; FEATURE:  
;; NAME/KEY: 3'UTR  
;; LOCATION: 6855..7175  
;; US-08-450-272-8

Query Match 3.8%; Score 49; DB 3; Length 7175;  
Best Local Similarity 48.0%; Pred. No. 0.0091;  
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;  
Qy 250 GACGTCGCGTGCAGGTGAACCCGCGCGACGCGCTCGGTGCAGTGTTCACCTCGGGCGC 309  
Db 2751 GAGGCCCGAAGCGGAGAGCGGGGAGCCCGGTGCCCCGGAGAGCGGCCCGCGCCGAC 2810  
Qy 310 CGCACGCTGCAGCTGCAGGTGCGAGCCAGCCCGACGCCGATCGGGTTCCTGTCAA 369  
Db 2811 CGCA-GCCACAGCAAGAGGCGCGGGGCC--CCCGAGGCGCGAGCGCGCGCGA 2867  
Qy 370 CCGCGTGCCACGCGCGCGCGGAGATCCCGCGGATCCTGGCAGACGTTAGCCCCGTTTC 429  
Db 2868 GGCCCCAGGCCCGGAGGGCGCGCGGACACCGCGCGCGCTCCCCGGAGAGCGCGGCC 2927  
Qy 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGAGGCGAGCAGACA 489  
Db 2928 GAGCGGGAGCCCCGACGCCACCGCGGCACCGGCACCAAGATCCGAGCAAGAGTGGCCC 2987  
Qy 490 CCCACGAGAGGAGGGGAGCCCGGCATCTTCGGGGACCCCGGGAACCGGAGCCGAGAGAG 549  
Db 2988 GGCGCCAGGGCGAGCGCGCGCGCGGACCGCGCGCGCCCCCGGAGCGGGGCCCGCGGAG 3047  
Qy 550 GTGGCCGCGGAAAGCGGTCCCCCAGCGCGGAGGAGGGGGGATGTTTCAGGCTGCA 609  
Db 3048 GCGGAGAGCGGGGAGGAGCGCGCGCGGACCCGGGGCCCGGCAAGAGGCGACGCTGCT 3107

Qy 610 GGGCAGGCGCGGTGGGAGCAGCAGCCACCACCGGAGGACCGGAAACAGTGTGGCGCGGAT 668  
Db 3108 CACGAGGCTGTGGAGAGGAGACCCACCGGAGAGGAGGCCACCGGAGAGGAGGCTGAGAT 3166  
Search completed: January 5, 2005, 19:22:37  
Job time : 137 secs



**THIS PAGE BLANK (USPTO)**



121 CGACCCGCGCCCTCTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTC 180  
 Db |||||  
 121 CGACCCGCGCCCTCTCTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTC 180  
 Qy |||||  
 181 GACAGCCACGAGCGGCGACAGCTATGGCCCTGCTGTCGCGATGGTCCCGGTGGTC 240  
 Db |||||  
 181 GACAGCCACGAGCGGCGACAGCTATGGCCCTGCTGTCGCGATGGTCCCGGTGGTC 240  
 Qy |||||  
 241 AGCAGCGTGAGCTGCGGTGAGTGAACCGCGCGCGAGCTCGGTGAGTGTCA 300  
 Db |||||  
 241 AGCAGCGTGAGCTGCGGTGAGTGAACCGCGCGCGAGCTCGGTGAGTGTCA 300  
 Qy |||||  
 301 CTCGGCGCGCGACGCTGCGAGCTTCAGGGTGCAGGACGCCCGCGATCGGGT 360  
 Db |||||  
 301 CTCGGCGCGCGACGCTGCGAGCTTCAGGGTGCAGGACGCCCGCGATCGGGT 360  
 Qy |||||  
 361 TCCTGTCAACCGGTGGCCACCGCGCGCGGAGATCCCGGATCTCGGACCGGTA 420  
 Db |||||  
 361 TCCTGTCAACCGGTGGCCACCGCGCGCGGAGATCCCGGATCTCGGACCGGTA 420  
 Qy |||||  
 421 GCCCGTCTCTCGGTGACCTTCGTGGCTCTCTCTCACTGGAGTTGCGGAGGC 480  
 Db |||||  
 421 GCCCGTCTCTCGGTGACCTTCGTGGCTCTCTCTCACTGGAGTTGCGGAGGC 480  
 Qy |||||  
 481 AGGACAGACCCACGAGGAGGAGGAGCCCGGATCTCGGGACCGGGAACCGGAG 540  
 Db |||||  
 481 AGGACAGACCCACGAGGAGGAGGAGCCCGGATCTCGGGACCGGGAACCGGAG 540  
 Qy |||||  
 541 CCGAGAGGTGGCCGCGAGGAAAGCGTCCCGGAGCGAGGAGGCGATGTT 600  
 Db |||||  
 541 CCGAGAGGTGGCCGCGAGGAAAGCGTCCCGGAGCGAGGAGGCGATGTT 600  
 Qy |||||  
 601 CAGGCTGACGAGGAGGAGGAGGAGCCCGGATCTCGGGACCGGGAACCGGAG 660  
 Db |||||  
 601 CAGGCTGACGAGGAGGAGGAGGAGCCCGGATCTCGGGACCGGGAACCGGAG 660  
 Qy |||||  
 661 GCGGCGATGAGTCTGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 Db |||||  
 661 GCGGCGATGAGTCTGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 Qy |||||  
 721 GACCCGCTGATTCGGATGCGCTCGAGACCGAGGAGGAGGAGGAGGAGGAGGAG 780  
 Db |||||  
 721 GACCCGCTGATTCGGATGCGCTCGAGACCGAGGAGGAGGAGGAGGAGGAGGAG 780  
 Qy |||||  
 781 AAGGAGCGCTGCTGCTTTCAGTTCCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGAC 840  
 Db |||||  
 781 AAGGAGCGCTGCTGCTTTCAGTTCCTTAGAGCAGAGTACGGCTACTATCACTGCAAGGAC 840  
 Qy |||||  
 841 TGCAAAATCCGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 Db |||||  
 841 TGCAAAATCCGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 Qy |||||  
 901 CTTCAAAAGTCTGCGGAGTGTGAGAAATCCTACAAACCTTTACAGAGTGGAGACAT 960  
 Db |||||  
 901 CTTCAAAAGTCTGCGGAGTGTGAGAAATCCTACAAACCTTTACAGAGTGGAGACAT 960  
 Qy |||||  
 961 CACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCGGAGTGGAGTGGAGTGGAGTGGAG 1020  
 Db |||||  
 961 CACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCGGAGTGGAGTGGAGTGGAGTGGAG 1020  
 Qy |||||  
 1021 CCCTTAAAGCGCCCATCGGAAAGTGTGAGGAGTGAAGGACAAACCGCTGTCTCTG 1080  
 Db |||||  
 1021 CCCTTAAAGCGCCCATCGGAAAGTGTGAGGAGTGAAGGACAAACCGCTGTCTCTG 1080  
 Qy |||||  
 1081 CGACAGACCTTACAGTCTCAAAATACATCATTTAGTGGAGTCAAAACGTTTCTGTAGA 1140  
 Db |||||  
 1081 CGACAGACCTTACAGTCTCAAAATACATCATTTAGTGGAGTCAAAACGTTTCTGTAGA 1140  
 Qy |||||  
 1141 TGGGGCTAATGGAATGACAGTGGAGTGTCTCCCTCTTCACTCTTCCCTTCCAAAT 1200  
 Db |||||  
 1141 TGGGGCTAATGGAATGACAGTGGAGTGTCTCCCTCTTCACTCTTCCCTTCCAAAT 1200

1201 TCTTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAAGGTATTGCAACA 1260  
 Db |||||  
 1201 TCTTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAAGGTATTGCAACA 1260  
 Qy |||||  
 1261 AAAAAAATAAAAAA 1277  
 Db |||||  
 1261 AAAAAAATAAAAAA 1277  
 Qy |||||  
 RESULT 2  
 US-10-475-502-1  
 ; Sequence 1, Application US/104/5502  
 ; Publication No. US20040254132A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Pei  
 ; APPLICANT: Wu, Xuemei  
 ; APPLICANT: Matzok, Martin M.  
 ; APPLICANT: Bai, Yuchen  
 ; TITLE OF INVENTION: Ovary-Specific Genes and Proteins  
 ; FILE REFERENCE: P01925U57  
 ; CURRENT APPLICATION NUMBER: US/10/475,502  
 ; CURRENT FILING DATE: 2003-10-22  
 ; PRIOR APPLICATION NUMBER: PCT US02/13245  
 ; PRIOR FILING DATE: 2002-04-26  
 ; PRIOR APPLICATION NUMBER: US 09/844,864  
 ; PRIOR FILING DATE: 2001-04-27  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1258  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; US-10-475-502-1  
 Query Match 96.6%; Score 1234.2; DB 18; Length 1258;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1250; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
 Qy 3 GGCGGGGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 62  
 Db 1 GGCGGGGAGGCGCGGAGCGACCCATGTTCCCGGCGAGCAGTTCACCCCTGCCGCA 60  
 Qy 63 TCCTTATCCGAGCGCCACAAAGCGGGGATGGGTGGAGTTCCGAGCGAGGCTGCCG 122  
 Db 61 TCCTTATCCGAGCGCCACAAAGCGGGGATGGGTGGAGTTCCGAGCGAGGCTGCCG 120  
 Qy 123 ACCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTGCA 182  
 Db 121 ACCGCGCCCTCTCTTCTCCCGGCTACAGACAGCTCATGGCGCGGAGTACGTGCA 180  
 Qy 183 CAGCCACAGCGGCGACAGCTCATGGCCCTCTGTGCGGATGGGTCCCGGTGGTCAG 242  
 Db 181 CCGCCACAGCGGCGACAGCTCATGGCCCTCTGTGCGGATGGGTCCCGGTGGTCAG 240  
 Qy 243 CAGCGGTACCGTGGGTGAGGTGAACCCCGCGCGAGCGCTCGGTGAGTGTCACT 302  
 Db 241 CAGCGGTACCGTGGGTGAGGTGAACCCCGCGCGAGCGCTCGGTGAGTGTCACT 300  
 Qy 303 CGGCGCGCGACGCTGACGCTTCAGGGTGCAGGCGCGAGCCCGGACCGCCGATCGGGTTC 362  
 Db 301 CGGCGCGCGACGCTGACGCTTCAGGGTGCAGGCGCGAGCGCCCGGATCGGGTTC 360  
 Qy 363 CTGTCAACCCCGTGGCCACCGCGCGCGGAGATCCCGGATCCCTGGAGCGGTAGC 422  
 Db 361 CTGTCAACCCCGTGGCCACCGCGCGCGGAGATCCCGGATCCCTGGAGCGGTAGC 420  
 Qy 423 CCGGTTCGTGCTGGTACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGAGGCGAG 482  
 Db 421 CCGGTTCGTGCTGGTACCTTCTGTGGCTCTCTCTCTCACTGGAGGTTGCGGAGGCGAG 480  
 Qy 483 GCAGACACCCACGAAAGGAGGAGGAGCCCGGATCTTCGGGAGCCCGGAGCCCGAGCC 542  
 Db 481 GCAGACACCCACGAAAGGAGGAGGAGCCCGGATCTTCGGGAGCCCGGAGCCCGAGCC 540

QY 543 GAGAGAGTGGCCGCGAGGAAGCGGTCCCCAGCCGCGAAGCGAGGAGGCGATGTTCA 602  
DB 541 GAGAGAGTGGCCGCGAGGAAGCGGTCCCCAGCCGCGAAGCGAGGAGGCGATGTTCA 600  
QY 603 GGTCTCAGGCGAGGCGCGGTGGAGCAGCAGCACCACCGAGGACCGGAACAGTGTGCG 662  
DB 601 GGTCTCAGGCGAGGCGCGGTGGAGCAGCAGCACCACCGAGGACCGGAACAGTGTGCG 660  
QY 663 GCGGATGAGTCTGAGCTGGGAGCGAGGCGCATGTCTCCGCGAGAGTGGCTCAGGA 722  
DB 661 GCGGATGAGTCTGAGCTGGGAGCGAGGCGCATGTCTCCGCGAGAGTGGCTCAGGA 720  
QY 723 CCCCGGTGATTCGATGCGCTTCGAGACCGAGCCCTCCCCGAAAGCAGGAGCAGCA 782  
DB 721 CCCCGGTGATTCGATGCGCTTCGAGACCGAGCCCTCCCCGAAAGCAGGAGCAGCA 780  
QY 783 GAGGCGCTCGGCTTTCCAGTTCTTAGAGCAGAACTAGAGTGTGCTGCTTCTTCACTG 842  
DB 781 GAGGCGCTCGGCTTTCCAGTTCTTAGAGCAGAACTAGAGTGTGCTGCTTCTTCACTG 840  
QY 843 CAATATCCGCTGGGAGCGGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902  
DB 841 CAATATCCGCTGGGAGCGGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899  
QY 903 TCAAAACAGTCTGCGAGTGTGTGAGAAATCTTACAACTTACAGAGTGGAGACATCA 962  
DB 900 TCAAAACAGTCTGCGAGTGTGTGAGAAATCTTACAACTTACAGAGTGGAGACATCA 959  
QY 963 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022  
DB 960 CCTGTCAAAGTTGTAAGAACTAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
QY 1023 CTAAACGCCCTCGGAGGAGT 1082  
DB 1020 CTAAACGCCCTCGGAGGAGT 1079  
QY 1083 ACAGACACCTTCAGCTTCAAATACATCATTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1142  
DB 1080 ACAGACACCTTCAGCTTCAAATACATCATTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1139  
QY 1143 GGGCTAAATGGAATGGAAGT 1202  
DB 1140 GGGCTAAATGGAATGGAAGT 1199  
QY 1203 TTCATGACAGACAGT 1261  
DB 1200 TTCATGACAGACAGT 1258

RESULT 3

US-09-844-864-18  
; Sequence 18, Application US/09844864  
; Patent No. US20020042926A1  
; GENERAL INFORMATION:  
; APPLICANT: Matzuk, Martin  
; APPLICANT: Ren, Yongsheng  
; APPLICANT: Wu, Xuemei  
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS  
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48  
; CURRENT APPLICATION NUMBER: US/09/844, 864  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/106, 020  
; PRIOR FILING DATE: 1998-10-28  
; PRIOR APPLICATION NUMBER: PCT/US99/25209  
; PRIOR FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 814  
; TYPE: DNA  
; ORGANISM: mus musculus  
US-09-844-864-18

Query Match 62.6%; Score 799.4; DB 9; Length 814;  
Best Local Similarity 99.9%; Pred No. 3e-227;  
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 GCGCGGCGAGGCGCGGAGCAGCACCATGTTTCCCGCGAGCAGCTTCCACCCCTGCGCGCA 62  
DB 1 GCGCGGCGAGGCGCGGAGCAGCACCATGTTTCCCGCGAGCAGCTTCCACCCCTGCGCGCA 60  
QY 63 TCCTTATCCGCGAGGCGCGGAGTGGCTTGGAGGTTCCGAGCAGGCGGCTGCGG 122  
DB 61 TCCTTATCCGCGAGGCGCGGAGTGGCTTGGAGGTTCCGAGCAGGCGGCTGCGG 120  
QY 123 ACCGCGCGCCCTCTCTTCTTCCCGGTACAGACAGCTCATGTGGCGCGGAGTAGCTCGA 182  
DB 121 ACCGCGCGCCCTCTCTTCTTCCCGGTACAGACAGCTCATGTGGCGCGGAGTAGCTCGA 180  
QY 183 CAGCCACAGCGGCGCAGCTCATGTGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 242  
DB 181 CAGCCACAGCGGCGCAGCTCATGTGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240  
QY 243 CAGCGGTGACGCTCGGCTGAGGTGAACCCCGCGCGGAGCGCTTCCGTCAGTGTTC 302  
DB 241 CAGCGGTGACGCTCGGCTGAGGTGAACCCCGCGCGGAGCGCTTCCGTCAGTGTTC 300  
QY 303 CGGCGCGCGCAGCTGACGCTGAGGTCGAGGTCGAGCCCGCGCGGAGCGCTTCCGTCAGTGTTC 362  
DB 301 CGGCGCGCGCAGCTGACGCTGAGGTCGAGGTCGAGCCCGCGCGGAGCGCTTCCGTCAGTGTTC 360  
QY 363 CTGTCAAACCCCGTGGCGCAGCGCGCGCGGAGATCTTCCGCGATCTTCCGCGAGACCGTAGC 422  
DB 361 CTGTCAAACCCCGTGGCGCAGCGCGCGCGGAGATCTTCCGCGATCTTCCGCGAGACCGTAGC 420  
QY 423 CCCGTTCTGTCTGTGACCTTCTGTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 482  
DB 421 CCCGTTCTGTCTGTGACCTTCTGTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480  
QY 483 GCAGACACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542  
DB 481 GCAGACACCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
QY 543 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 602  
DB 541 GAGAGAGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
QY 603 GGTCTGAGGCGAGGCGCGGTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 662  
DB 601 GGTCTGAGGCGAGGCGCGGTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 660  
QY 663 GCGGATGAGTCTGAGCTGGGAGCGAGGAGCGCATGTCTCCGCGAGAGTGGCTCAGGA 722  
DB 661 GCGGATGAGTCTGAGCTGGGAGCGAGGAGCGCATGTCTCCGCGAGAGTGGCTCAGGA 720  
QY 723 CCCCGGTGATTCGATGCGCTTCGAGACCGAGCCCTTCCCGCGAAAGCAGGAGCAGGAGCA 782  
DB 721 CCCCGGTGATTCGATGCGCTTCGAGACCGAGCCCTTCCCGCGAAAGCAGGAGCAGGAGCA 780  
QY 783 GGAGGCGCTGCGTTTCCAGTT 803  
DB 781 GGAGGCGCTGCGTTTCCAGTT 801

RESULT 4

US-10-475-502-11  
; Sequence 11, Application US/10475502  
; Publication No. US20040254132A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Pei  
; APPLICANT: Wu, Xuemei  
; APPLICANT: Matzuk, Martin M.  
; APPLICANT: Bai, Yuchen  
; TITLE OF INVENTION: Ovary-Specific Genes and Proteins  
; FILE REFERENCE: P01925US7

```
; CURRENT APPLICATION NUMBER: US/10/475,502
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: PCT US02/13245
; BEST FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 09/844,864
; FIRST FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 6873
; TYPE: DNA
; ORGANISM: mus musculus
US-10-475-502-11

Query Match      62.6%; Score 799.4; DB 18; Length 6873;
Best Local Similarity 99.9%; Pred. No. 7.4e-227;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY    3   GGCGGGGAGCGCGGGAGCCACCAATGTTCCTCCGGGAGCAGCACGTTCACCCCTGCCCGCA 62
DB    2513  GGCGGGGAGCGCGGGAGCGACCAATGTTCCTCCGGGAGCAGCATTTCCACCCCTGCCCGCA 2572

QY    63   TCCTTATCCAGGCACCACAAAGCCGGGATGCTCGAAGTTCGAGCCAGGGGCTGCCG 122
DB    2573  TCCTTATCCAGGCACCACAAAGCCGGGATGCTCGAAGTTCGAGCCAGGGGCTGCCG 2632

QY    123  ACCCGCGCCCCCTCTCTCTCCCGGTACAGACAGCTCATCGCCGCGGAGTACGTCGA 182
DB    2633  ACCCGCGCCCCCTCTCTCTCCCGGTACAGACAGCTCATCGCCGCGGAGTACGTCGA 2692

QY    183  CAGCCACAGCGGGCACAGCTCATGCGCCCTGCTGTGCGGATGGGTCCC CGGTCGTCAG 242
DB    2693  CAGCCACAGCGGGCACAGCTCATGCGCCCTGCTGTGCGGATGGGTCCC CGGTCGTCAG 2752

QY    243  CAGCCGTGACGCTCGGTGACAGTGTAACCGCGCGCGGAGCCCTCGGTGACAGTGTCACT 302
DB    2753  CAGCCGTGACGCTCGGTGACAGTGTAACCGCGCGCGGAGCCCTCGGTGACAGTGTCACT 2812

QY    303  CGGCGCGGCACGCTGACGCTGACAGGCTGCGAGCGCCGAGCCCGACGCCCGATCGGGTTC 362
DB    2813  CGGCGCGGCACGCTGACGCTGACAGGCTGCGAGCCCGACGCCCGATCGGGTTC 2872

QY    363  CTGTCAAACCCGTGGCCACGCGCGCGCGGGAGATCCCGCGATCTCGGCAGACCGTAGC 422
DB    2873  CTGTCAAACCCGTGGCCACGCGCGCGCGGGAGATCCCGCGATCTCGGCAGACCGTAGC 2932

QY    423  CCCGTTCTCGTCGCTGACCTTCTGTGCGCTCTCTCTCTCTACTGAGATTCCGGAGGCGAG 482
DB    2933  CCCGTTCTCGTCGCTGACCTTCTGTGCGCTCTCTCTCTCTACTGAGATTCCGGAGGCGAG 2992

QY    483  GCAGACACCCACGAAGGAGAGGGAGCCCGGCATCTTCGGGACCCCGGAAACCGAGCC 542
DB    2993  GCAGACACCCACGAAGGAGAGGGAGCCCGGCATCTTCGGGACCCCGGAAACCGAGCC 3052

QY    543  GAGAGAGTGGCCGCGAGGAAGCGGTTCCTCCAGCGCGAGCGAGAGGGCGCATGTTCA 602
DB    3053  GAGAGAGTGGCCGCGAGGAAGCGGTTCCTCCAGCGCGAGCGAGAGGGCGCATGTTCA 3112

QY    603  GGCTGCAGGGCAGGCCCGGTTGGGAGCAGCAGCCACCCGAGAGCCCGAACAGTGTGGC 662
DB    3113  GGCTGCAGGGCAGGCCCGGTTGGGAGCAGCAGCCACCCGAGAGCCCGAACAGTGTGGC 3172

QY    663  GCGATGACGTCAGCCTGGGAGCAGAGCCATGTCTCTCGCAGAGATGCTCAGGA 722
DB    3173  GCGATGACGTCAGCCTGGGAGCAGAGCCATGTCTCTCGCAGAGATGCTCAGGA 3232

QY    723  CCCCGGTGATTTCGATGCCCTTCAGACACAGGCTCCCGCAAGCAGCAGCAGACAA 782
DB    3233  CCCCGGTGATTTCGATGCCCTTCAGACACAGGCTCCCGCAAGCAGCAGCAGACAA 3292

QY    783  GGAGCGCTCGCTTTCCAGTT 803
DB    3293  GGAGCGCTCGCTTTCCAGTT 3313
```









Qy 209 CCTGCTCTCGGGATGGGTCCCCGGTCTGCTGAGAGCCGTGAGAGCTGCGGTGAGGTGA 268  
Db 464 GCGCGGGCGGGCGGGTTCGGCGCCCGGGCCCCCGCCCCCGCGCGCGCGCGCGCGCG 405  
Qy 269 ACCCGCGCGCGACG-----CCTCGGTGAGTGTTCACCTCGGCGCGCGACGCTGCAGC 322  
Db 404 GCGCGGTTCGGCGCGGGCG 345  
Qy 323 CTGAGGGTTCGCGAGCCCGACGCCCCGATCGGGTTCCTGTCAACCCCGTGGCCACG 382  
Db 344 GGGCGGGGGCG 285  
Qy 383 CCGCGCGCGGGAGATCCCGGATCCTCGGAGACCGTAGACCGCTTCGTCTCGTGACCT 442  
Db 284 GGTGCGCGCGCGGGCGCGCGCGCGCTGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 225  
Qy 443 TCTGTGGCTCTCTCTCTCACTGGAGGTTGGGGAGGAGCAGACACCCAGAGGGAG 502  
Db 224 GGTGCGCGCGCGCGCGCGCGCGGTGGTGGCGGGGGGGGGCGCGCGCGCGCGCGCGCG 165  
Qy 503 AGGGAGAGCCCGCATCTCTCGGGACCCGGGAACCGGAGCCGAGAGAGGTGGCGCGAGGA 562  
Db 164 GCG 105  
Qy 563 AAGCGGTCCCGCAGCGCGAGAGCGAGGAGGCGATGTTGAGGCTGAGGGCAGCGCGGT 622  
Db 104 CCGGGCG 45  
Qy 623 GGGAGCAGCAGCCACAC 640  
Db 44 GCG 27

RESULT 15  
US-10-425-115-17653/c  
; Sequence 17653, Application US/10425115  
; Publication No. US2004021427A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 17653  
; LENGTH: 972  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(972)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRF4577\_116102C.1  
US-10-425-115-17653

Query Match 4.3%; Score 55; DB 18; Length 972;  
Best Local Similarity 46.6%; Pred. No. 1.6e-05;  
Matches 269; Conservative 0; Mismatches 304; Indels 4; Gaps 3;  
Qy 21 CGCACCATGTTCCCGCGGAGCAGTTCACCCCTGCGCGCATCTTATCCGAGGCCAC 80  
Db 819 CACCCCG 760  
Qy 81 CAAAGCGGGATGGTGGAGGTTGCGAGCAGGGGCTGCCGACCGCGCGCGCGCGCGCGCTT 140  
Db 759 CACCCCG 700

Qy 141 CCTCCCGGCTACAGACAGCTCATGCGCGGAGTACGTG--ACAGCCACACAGCGGGCA 198  
Db 699 GCG 640  
Qy 199 CAGTCTATGCGCCTGCTGTCGCGGATGGGTCCCGGTTCGGTTCAGCAGCCGTGACGTGCG 258  
Db 639 CCGCCCG 580  
Qy 259 GTGAGGTGAAC--CGCT 317  
Db 579 GCG 520  
Qy 318 GCAGCCTGACGGGTGCGAGCCCGAGCCCGGATCGGGTTCCTGTCAACCCCGGTGG 377  
Db 519 GGGCG 460  
Qy 378 CCAAGCGCGCGCGCGGAGATCCCGCGGATCCTGCGAGACCGTAGCCCGCTTCTGTCCTGT 437  
Db 459 CCG 400  
Qy 438 GACCTTCTGTGGCTCTCTCTCACTGAGGTTGCGGGAGGCGAGCAGACACCCAGAA 497  
Db 399 CCG 340  
Qy 498 GGGAGAGGGGAGCGCGCATCTCTCGGGGACCCCGGAAACCGGAGCCGAGAGGTGGCGCG 557  
Db 339 GGAG--CGCGGGCG 281  
Qy 558 GAGGAAAGCGGTCTCCCGCAGCGCGCGAGCAGGAGGGGC 594  
Db 280 GGGCG 244

Search completed: January 5, 2005, 21:16:57  
Job time : 756 secs